ANALYSIS OF TRANSCRIPTIONAL REGULATION OF X-LINKED GENES ON THE ACTIVE AND INACTIVE X CHROMOSOMES

By

MICHAEL DUANE LITT

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE
UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1997

I would like to dedicate this dissertation to my beautiful wife Christine for her help and support throughout the years.

ACKNOWLEDGEMENTS

I would like to acknowledge my mentor, Thomas P. Yang, whose guidance, support and faith in my abilities were important in my development as a person and scientist. I would also like to thank my wife who supported me throughout school both emotionally and scholastically. Furthermore, I would like to acknowledge Ian Hornstra who taught me in vivo footprinting and other members of Dr. Yang's laboratory for their help.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.. iii

LIST OF FIGURES... vii

ABBREVIATIONS.. x

ABSTRACT.. xii

CHAPTERS

1. INTRODUCTION.. 1
    X Inactivation.. 1
    The Role of DNA-Binding Regulatory Proteins......................... 5
    The Role of DNA Methylation... 6
    The Role of Chromatin Structure.. 11
    X Chromosome Inactivation in Males.................................. 12
    HPRT Gene... 14
    The Human XIST Gene.. 15
    Specific Aims and Goals.. 16
2. IN VIVO FOOTPRINTING AND HIGH-RESOLUTION METHYLATION ANALYSIS OF THE MOUSE HPYOXANTHINE PHOSPHORIBOSYLTRANSFERASE GENE 5' REGION ON THE ACTIVE AND INACTIVE X CHROMOSOME.. 19
    Introduction.. 19
    Materials and Methods... 23
        Cell Lines... 23
        Dimethyl Sulfate (DMS) Treatment.............................. 23
        Separation and Isolation of the Active and Inactive 5' Regions of the *M. musculus* HPRT Gene.. 24
        Hydrazine Treatment for Genomic Sequencing.................... 25
        Preparation of the Control DNA Sequencing Ladder............. 26
        Ligation-Mediated PCR.. 26
        Gel Electrophoresis and Electrotransfer........................ 27
        Probe Synthesis, Hybridization, and Washing.................... 27
    Results... 29
        In Vivo Footprinting... 29
        Genomic Sequencing... 40
    Discussion... 47

In Vivo Footprinting.. 47
Genomic Sequencing.. 49
Implications for X Chromosome Inactivation.................. 52
3. ELECTROPHORETIC GEL MOBILITY SHIFT ASSAYS OF A DNA-PROTEIN INTERACTION SPECIFIC TO THE ACTIVE ALLELE OF THE HUMAN AND MOUSE HPRT GENE... 54
Introduction.. 54
Material and Methods.. 56
Nuclear Extracts... 56
Preparation of Double Stranded Oligos for Gel Mobility-Shift Assays.. 56
Preparation of Sequence-Specific Competitors................. 57
Electrophoretic Gel Mobility-Shift Assays..................... 59
Results... 60
Discussion.. 65
4. 5-AZADEOXYCYTIDINE-INDUCED CHROMATIN REMODELING OF THE INACTIVE X-LINKED HPRT GENE PROMOTER OCCURS PRIOR TO TRANSCRIPTION FACTOR BINDING AND GENE REACTIVATION..................... 69
Introduction... 69
Materials and Methods.. 72
Cell Lines... 72
5-Aza-2`-deoxycytidine (5aCdr) Treatment.................... 72
MspI Treatment of Chromatin.................................. 73
RT-PCR of Human HPRT mRNA.............................. 75
Prepartion of DNA: In Vitro Dimethyl Sulfate (DMS) Treatment and DNA Isolation................................. 76
Ligation-Mediated PCR (LMPCR) and Detection of In Vivo Footprints.................................. 76
Quantification of Footprints by Autoradiogram Densitometry.. 77
Results.. 78
Changes in Chromatin Structure of the 5' Region During 5aCdr-induced Reactivation of the HPRT Gene.. 78
Appearance of Human HPRT mRNA During 5aCdr-Induced Reactivation of the HPRT Gene.. 80
Binding of Transcription Factors During 5aCdr-Induced Reactivation of the HPRT Gene............... 80
Discussion.. 89
5. TRANSCRIPTIONAL INACTIVATION OF THE MOUSE HPRT GENE DURING SPERMATOGENESIS.. 92
Introduction.. 92
Materials and Methods.. 93

Separation of Spermatogenic Cells................................ 93
DMS Treatment.. 94
Ligation-Mediated PCR... 94
Quantification of Footprints by Autoradiogram Densitometry.. 94
Results... 95
Disscussion.. 104
6. IN VIVO FOOTPRINTING OF THE MINIMAL HUMAN XIST PROMOTER ON THE ACTIVE AND INACTIVE X CHROMOSOMES.. 106
Introduction.. 106
Materials and Methods.. 108
Cell Lines.. 108
Dimethyl Sulfate Treatment and DNA Isolation................ 109
Ligation-Mediated PCR (LMPCR) and Detection of In Vivo Footprints.................................. 109
Results... 110
In Vivo Footprinting of the Human XIST Promoter............ 110
In Vivo Footprinting of the Mutant XIST Promoter............ 114
Discussion... 122
In Vivo Footprinting.. 122
Implications for X Inactivation.................................. 126
7. CONCLUSIONS AND FUTURE DIRECTIONS........................... 127
Implications for X Inactivation.. 131

REFERENCES LIST.. 154

BIOGRAPHICAL SKETCH.. 169

LIST OF FIGURES

Figure

2.1 - Location of LMPCR primers for analysis of the mouse HPRT 5' region... 30

2.2 - In vivo footprint analysis of the upper strand using primer set U... 33

2.3 - In vivo footprint analysis of the lower strand using primer set J.. 35

2.4 - Summary of in vivo footprints in the 5' region of the active *M. musculus* HPRT gene... 37

2.5 - Alignment of the 5' region from the mouse and human HPRT genes.. 39

2.6 - Genomic sequencing of the upper strand and lower strand using primer sets J and U... 43

2.7 - Summary of the methylation pattern from the 5' region of the inactive mouse HPRT gene... 46

3.1 - Diagram of the oligos and sequence-specific competitors used in gel mobility shift assays... 58

3.2 - Electrophoretic mobility-shift assays using the 195/196, 197/198, 201/202, and 216/217 double stranded oligos as probes.................. 61

4.1 - Summary diagram of footprints, PCR primers, probes, and DNA fragments used to analyze the human HPRT gene.. 74

4.2 - Nuclease sensitivity of the 5' CpG island of the human HPRT gene following treatment of X8-6T2 cells with 5aCdr.. 79

4.3 - RT-PCR analysis of human HPRT mRNA in X8-6T2 cells following treatment with 5aCdr....................................... 81

4.4 - In vivo footprint analysis at position -91 following treatment of X8-6T2 cells with 5aCdr..................................... 83

4.5 - In vivo footprint analysis of the GC box region following treatment of X8-6T2 cells with 5aCdr.. 86

4.6 - Time course of HPRT mRNA levels, nuclease sensitivity of chromatin, and transcription factor binding at positions -91, -198, & -210 following 5aCdr treatment.. 88

5.1 - Location of LMPCR primers for analysis of the mouse HPRT 5' region... 96

5.2 - DMS in vivo footprinting of the upper strand of the mouse HPRT gene at different stages of spermatogenesis.. 99

5.3 - Densitometry of four footprints on the upper strand at different stages of spermatogenesis..................................... 101

5.4 - DMS in vivo footprinting of the lower strand of the mouse HPRT gene at different stages of spermatogenesis... 102

5.5 - Densitometry of the +14 footprint on the lower strand at different stages of spermatogenesis.. 103

6.1 - Location of LMPCR primers for analysis of the human XIST 5' region... 111

6.2 - In vivo footprint analysis of the lower strand of the normal XIST promoter using primer set X........................... 113

6.3 - In vivo footprint analysis of the upper strand of the normal XIST promoter using primer set Y........................... 116

6.4 - In vivo footprint analysis of the lower strand of the mutant XIST promoter using primer set X........................... 119

6.5 - In vivo footprint analysis of the lower strand of the mutant XIST promoter using primer set X........................... 121

6.6 - Summary of in vivo footprints in the normal and mutant XIST promoter on the inactive X Chromosome.. 124

7.1 - Comparison of in vivo footprints from the active allele of 3 X-linked genes.. 133

7.2 - The methylation patterns of the promoters of four X-linked housekeeping genes... 134

7.3 - Model for the transcriptional regulation of X-linked genes by chromatin domains.. 136

ABBREVIATIONS

3’	Three prime
5aCdr	5-aza-2’-deoxycytidine
5-azaC	5-azacytidine
5’	Five prime
A	Adenine
ATP	Adenosine Triphosphate
bp	Base pair
C	Cytosine
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytosine triphosphate
dGTP	Deoxyguanosine triphosphate
dNTP	Deoxyribonucleotide triphosphate
dTTP	Deoxythymidine triphosphate
D-MEM	Dulbecco’s Modified Eagle Medium
DMS	Dimethyl sulfate
Dnmt	Methyltransferase mutant
DNA	Deoxyribonucleic acid
DNase I	Deoxyribonuclease I
EDTA	Ethylenediamine tetra-acetic acid
ES	Embryonic stem cells
FBS	Fetal bovine serum
g	One force of gravity
G	Guanine
GMP	Guanosine monophosphate
G6PD	Glucose-6-phosphate dehydrogenase
HAT	Hypoxanthine, aminopterin, thymidine
HPRT	Hypoxanthine phosphoribosyltransferase
IMP	Inosine monophosphate
kb	Kilobase
LMPCR	Ligation-mediated polymerase chain reaction
ml	Milliliter
mM	Millimolar
M	Molar
mRNA	Messenger ribonucleic acid
Pdha-1	Pyruvate dehydrogenase
Phka	Phosphorylase kinase
PBS	Phosphate buffered saline
PGK	Phosphoglycerate kinase
PCI	Phenol, chloroform, isoamyl alcohol
PCR	Polymerase chain reaction
P/S	Penicillin-streptomycin
pmol	Picomole
RNA	Ribonucleic acid
SDS	Sodium dodecylsulfate

T	Thymidine
TBE	Tris, boric acid, EDTA
TBP	TATA binding protein
TE	10 mM Tris, 1 mM EDTA
Tris	Hydroxymethyl aminomethane
ug	Microgram
ul	Microliter
uM	Micromolar
Xce	Mouse X chromosome controlling element
XIC	X inactivation center
XIST	Human X inactive specific transcript gene
Xist	Mouse X inactive specific transcript gene
Zfx	Zinc finger x gene

Abstract of Dissertation Presented to the Graduate School of the University of Florida in
Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

ANALYSIS OF TRANSCRIPTIONAL REGULATION OF X-LINKED GENES ON THE ACTIVE AND INACTIVE X CHROMOSOMES

By

Michael D. Litt

August 1997

Chairperson: Thomas P. Yang
Major Department: Biochemistry and Molecular Biology

To obtain insight into the mechanisms involved in X chromosome inactivation, I have examined the binding of transcription factors to mouse hypoxanthine phosphoribosyltransferase (HPRT) gene on both the active and inactive X chromosomes, identified the methylation patterns of both the active and inactive mouse HPRT promoters, charted the binding of factors to the human HPRT gene during reactivation by 5-aza-2'-deoxycytidine (5aCdr), examined the timing of X inactivation during mouse spermatogenesis, and examined the binding of transcription factors to both a normal and mutant XIST promoter. In vivo footprinting analysis of the HPRT promoter revealed four footprinted regions on the active allele and none on the inactive allele. One footprint occurred over a novel 9 base pair sequence identical in both the human and mouse HPRT gene. Gel mobility shift analysis of this sequence revealed multiple DNA-protein complexes. Genomic sequencing indicated that the active HPRT allele was unmethylated while the majority of CpG's dinucleotides on the inactive allele were methylated. No obvious correlation between the methylation pattern and transcription factor binding was observed. During (5aCdr)-induced reactivation of the human HPRT gene on the inactive

X chromosome, acquisition of a nuclease sensitive chromatin conformation in the 5' region occurs prior to the appearance of HPRT mRNA. In vivo footprinting experiments indicated that this change in chromatin structure precedes the binding of factors suggesting that inactive X-linked genes were silenced by a repressive chromatin structure that prevents factor binding. In vivo footprinting analysis suggested that transcriptional silencing of the male X chromosome is largely complete just prior to or during the late pachytene stage of mouse spermatogenesis. Examination of the normal XIST allele and a mutant allele, which contains a rare mutation that results in skewed X inactivation in cis, revealed multiple footprints on both normal and mutant active alleles but no footprints on either the normal or mutant inactive alleles. However, over the mutated site different footprints were observed between the normal and mutant XIST alleles. This suggests the region containing the mutation may be involved in determining which X chromosome is silenced by XIST.

CHAPTER 1
INTRODUCTION

X Inactivation

In eutherian mammals, females have two X chromosomes and males have one X chromosome. This genotype results in a dosage imbalance of X-linked genes in male cells relative to female cells. As a mechanism that compensates for this imbalance of X-linked genes, one X chromosome in each female somatic cell is randomly selected during embryogenesis to be transcriptionally silenced or inactivated (90). Thus for most X-linked genes, both a transcriptionally active allele and transcriptionally inactive allele reside within each female somatic cell nucleus.

Evidence indicates that the process of X inactivation begins early in female embryogenesis with the majority of X-linked genes becoming transcriptionally silent on one X chromosome by the end of the late blastocyst stage in mouse female embryos (141). Prior to the early blastocyst stage, isoenzyme studies detect gene products from both X chromosomes. At the 40-50 cell blastocyst stage of the female mouse embryo, these studies indicate many X-linked genes are becoming inactivated (141). Therefore, one of the two X chromosomes is inactivated during the blastocyst stage. In the rare cases where the female cells contain more than two X chromosomes in a normal autosomal complement only one X chromosome remains active; the remaining X chromosomes are inactivated.

At least three steps are thought to be involved in X inactivation: initiation, spreading, and maintenance. Very little is known about the molecular mechanisms at each of these steps. However, current studies suggest two genetic loci are involved in X inactivation, the X-inactivation center (XIC) and the X chromosome controlling element (Xce). The XIC is a region on both the mouse and human X chromosomes which is required in cis for initiation of X inactivation to occur (103). The XIC was discovered by analysis of X:autosome translocations. In these studies the derivative X chromosomes lacking the XIC were unable to be inactivated (26, 33, 142). Similar translocation studies mapped the XIC to chromosomal band Xq13.2 in humans (16) and chromosomal band XD in mice (124). The human XIC has been localized to approximately a 1 megabase region (16) and a recent study by Lee et al. argues that the mouse Xic is contained in a 450 kb region (80).

In 1991 Brown et al. (14) discovered a gene within the human XIC with several unusual properties that made it a excellent candidate for a role in X inactivation. This gene, the XIST gene (X inactive specific transcript), was also discovered in the mouse Xic (12). The XIST gene is the only gene to date which is expressed exclusively from the inactive X chromosome (14). The onset of XIST expression precedes X inactivation (69). Thus, its abnormal expression is unlikely to be a consequence of X inactivation. XIST transcribes a noncoding RNA that remains associated with the inactive X chromosome in female interphase nuclei (15). Mutational studies of the XIST gene which generated nonfunctional XIST alleles in embryonic stem cells indicate that a functional XIST allele is necessary for X inactivation. In these studies only the X chromosome containing a functional XIST allele undergoes inactivation (118). However, cells carrying the targeted mutation were still capable of doing other functions hypothesized to occur at the onset of inactivation such as counting the number of X chromosomes present or randomly selecting which X was to be inactive. This suggests that the sequences responsible for counting and selection are separate from the XIST gene. Another study showed that 450 kb transgene containing the Xic/Xist is sufficient for chromosomal counting, choosing and initiation of X inactivation (80). When this 450 kb transgene was translocated into an autosome, the autosome was counted as another X chromosome, it was randomly selected like another X chromosome, and if it was selected for inactivation, it expressed the Xist gene. In addition, the Xist transcript painted the autosome into which it was translocated (80). This evidence suggests that the XIC and XIST play important roles in X inactivation and its initiation.

The X chromosome controlling element (Xce) (27) has been mapped to an interval with the Xic in mice. To date an XCE has not been described in humans. The alleles of this element appear to influence the randomness of X chromosome inactivation (22, 28). Four alleles of the Xce have been described Xce^a, Xce^b, Xce^c, and Xce^d (23, 25-27, 67). An X chromosome carry the Xce^a allele is more likely to be inactive than a chromosome carrying the Xce^b, Xce^c or Xce^d allele. The Xce does not appear to exert its effect by a cell lethality selection mechanism but rather by biasing the selection of which chromosome is to be active or inactive (67). Little is known about the structure or the molecular mechanisms of the Xce locus. Unlike the XIC/Xic, no genes have been mapped to the Xce locus. However, it does appear to play a role in X inactivation in mice.

Once the process of X inactivation is initiated, the silencing signal appears to propagate bi-directionally in a cis-limited manner toward the ends of the chromosome. This process is most evident in certain X:autosome translocations where heterochromatinization and inactivation have spread to a variable and limited extent into the attached autosomal material (123). Although the majority of X-linked genes are inactivated, a few genes, located mainly on the pseudo autosomal regions appear to escape inactivation (18, 42, 104, 106). Many of the genes that escape inactivation in humans do not escape inactivation in the mouse (1, 70, 157). At least one gene, Smcx, does escape inactivation in the mouse (2, 154). Recent evidence by Disteche (unpublished data) suggests that genes which escape inactivation are not skipped over for silencing during the spreading of inactivation. Instead, these genes are inactivated and then reactivated. Similar results of inactivation followed by reactivation were observed in the abnormal silencing of autosomal genes in X:autosome translocation (24). This suggests that initially all the genes on the inactive X chromosome undergo transcriptional silencing and that later genes which are missing an element which maintains their repressed state become reactivated. Presently, little is known about the molecular basis of the spreading of inactivation.

After one X chromosome has been selected for inactivation, the same X chromosome is maintained in an inactivate state throughout subsequent mitosis in all somatic cell progeny (46). As a result, in each female somatic cell nucleus an active and inactive copy of most X-linked genes is present. The molecular basis for the maintenance of differential expression of genes on the active versus inactive X chromosome is unclear. However, differential binding of regulatory proteins (46, 94), DNA methylation (46, 86, 107), DNA replication (45, 46), and chromatin structure (72, 112, 119) have been implicated.

The Role of DNA-Binding Regulatory Proteins

Since specific DNA sequences and regulatory proteins appear to have an essential role in many systems of transcriptional regulation (91, 105), DNA-binding regulatory proteins are likely to be involved in the process of X inactivation. One model proposed by McBurney (94) postulates that each X-linked gene, normally subject to inactivation, is regulated by the differential binding of repressor and activator proteins that interact with a cis-acting regulatory element(s) near or within the promoter of the gene. This suggests the coordinate regulation of X-linked genes by X inactivation is controlled by interactions between regulatory proteins and the cis-acting regulatory element(s). To date, a conserved regulatory sequence common to X-linked genes subject to X inactivation has not been identified by DNA sequence analysis. However, due to the potential degeneracy in the sequence specificity of DNA-binding regulatory proteins (11), the identification of a common X-linked DNA sequence, which interacts with a repressor or activator, will probably require functional binding studies.

Presently, the differential binding of sequence-specific DNA-binding proteins to the active and inactive X chromosomes has been examined only for two genes, the human HPRT and PGK-1 genes (63, 119, 122). In vivo footprinting analysis of the human HPRT and PGK-1 promoter revealed differential binding of transcription factors to their active and inactive alleles. In both genes, sequence-specific DNA-protein interactions are observed exclusively on the transcriptionally active promoter regions. No sequence-specific DNA-protein interactions are identified within the promoter of the inactive PGK-

1 and HPRT genes. Sequence analysis of the multiple footprints observed on the active HPRT and PGK-1 promoters indicated that the majority of the binding sites were for ubiquitous transcription factors such as Sp1 or NF-kB. Besides potential Sp1 binding sites, no footprinted sites contained unique sequences common to both genes. These studies suggest that the differential binding of regulatory proteins appears to play a role in the maintenance of the differential transcriptional regulation between the active and inactive X chromosomes. The absence of sequence-specific DNA-protein interactions on the inactive PGK-1 and HPRT alleles argues against a sequence-specific repressor protein binding to each X-linked gene. The fact that no unique DNA-binding regulatory factors or cis-acting sequences that are specific to an X-linked gene was observed in the PGK-1 and HPRT promoters suggests that the overall process of X inactivation is not regulated by specific DNA sequences in the promoter of each X-linked gene. Therefore, the fundamental regulation of X-linked genes by X inactivation may be occurring at another level (such as chromatin structure or DNA methylation).

To gain further insight into the role of the differential binding of DNA-binding regulatory proteins to the active and inactive alleles of X-linked genes and their influence on the mechanisms of X inactivation, in this dissertation I have employed in vivo footprinting to examine the differential binding of regulatory factors to the active and inactive 5' regions of the mouse HPRT gene. These data have been carefully compared to previous studies of the human HPRT and PGK-1 genes.

The Role of DNA Methylation

A growing body of evidence supports DNA methylation as a mechanism involved in maintaining the differential expression of housekeeping genes on the active and inactive X chromosome (61, 122, 143, 144). The methylation patterns of the active and inactive alleles of several X-linked genes has been examined using methyl-sensitive restriction enzymes (71, 89, 144, 156). In the X-linked PGK-1, glucose 6-phosphate dehydrogenase (G6PD) and HPRT genes, Hpa II and Hha I genomic digests indicated that the 5' region was hypomethylated at Hpa II and Hha I sites on the active X chromosome relative to the inactive X chromosome (71, 89, 144, 151). Therefore, for housekeeping genes there appears to be a correlation between hypermethylation of the inactive promoter and transcriptional silencing. This correlation does not appear to be true for non-housekeeping genes because methylation studies on the X-linked factor IX gene, a tissue-specific gene without a 5' GC island, showed no difference in the levels of DNA methylation on the active and inactive alleles (36). Although these studies suggest there is a correlation between transcriptional silence and DNA methylation of inactive X-linked housekeeping genes, they do not clearly identify a mechanism.

Since methyl sensitive enzyme studies fail to examine the majority of potential CpG DNA methylation sites, it is possible individual methylation sites missed by methyl sensitive enzymes may be directly involved in the differential transcriptional regulation of individual X-linked genes. Therefore, genomic sequencing studies which examine every CpG sites have been used to examine the high-resolution methylation pattern on the active and inactive promoters of three genes, the human and mouse PGK-1 (122, 143) and the human HPRT genes (61). The high-resolution methylation pattern of the human PGK-1 and HPRT genes indicated that CpG's are unmethylated in the 5' GC island on the active X chromosome and heavily methylated on the inactive X chromosome. Like the human PGK-1 and human HPRT gene all the CpG's are unmethylated on the active X chromosome in the mouse PGK-1 gene. Unlike these genes only one CpG is fully methylated on the inactive mouse PGK-1 promoter. Even though the correlation of hypermethylation of the inactive promoter relative to the active promoter remained true for these three X-linked genes, the absence of a common identifiable pattern of methylation on the inactive promoters among these three genes left no clear evidence for the mechanisms of action of DNA methylation.

Studies using 5-azacytidine (5-azaC), a DNA demethylating agent, have also implicated DNA methylation in the regulation of expression of X-linked genes (52, 53, 107, 144, 148). In hybrid cell lines carrying a single inactive X chromosome, treatment with 5-azacytidine caused the derepression of previously inactivated genes (including HPRT, PGK-1, G6PD, α-galactosidase). Genomic sequencing and methyl-sensitive enzyme studies indicate that the methylation pattern of reactivated genes is identical to the methylation pattern of the active allele (61, 122). This suggests that DNA methylation of the inactive allele is necessary to maintain transcriptional silence. However, it does indicate a clear molecular mechanism of action. In another experiment, Pfiefer et al. (120) examined inactive PGK-1 genes which had been treated with the demethylating agent 5-azacytidine in order to determine if there are any DNA methylation sites that are necessary to maintain transcriptional silencing. In this study, the CpG sites around the transcription initiation site remained methylated in all samples; the remaining sites showed variable degrees of methylation. Consequently, only a few sites of DNA methylation, primarily around the transcription initiation site, may be necessary to maintain a transcriptionally silent X-linked gene. Similar studies on other genes are necessary to confirm this observation.

Additional evidence for a role of DNA methylation in X inactivation has been provided by DNA transformation experiments (29, 86, 147, 148). In these experiments, purified-DNA containing either the active or inactive X chromosome was transfected into cells. DNA from the inactive X chromosome was significantly less efficient at transforming the cells than DNA from the active X chromosome (29, 86, 147, 148). Also, DNA isolated from a 5-azaC reactivated inactive X chromosome was significantly more efficient at transforming the cells than DNA form a non-reactivated inactive X chromosome (148). These results suggest that a property of the DNA itself was responsible for the differences in transformation efficiency. The most probable property responsible for the differences in transformation is the methylation state of the DNA from the X-linked genes examined.

In a recent study, Panning et al. (113) examined the expression of XIST in DNA methyltransferase mutant (Dnmt) embryos and embryonic stem (ES) cells. These mutants are unable to methylate their DNA. In wildtype, undifferentiated embryos and embryonic stem cells, Xist expression was detected at low levels from both female X chromosomes. The same level of Xist expression was observed for Dnmt in undifferentiated mutant ES cells. This suggests that DNA methylation does not regulate Xist expression in undifferentiated ES cells. Upon differentiation Xist expression was detected only from the inactive X chromosome in female cells. Differentiated Dnmt mutant ES cells had an unusual Xist expression pattern in which Xist was expressed and colocalized to both X chromosomes. In addition, X-linked gene expression was not detected from X chromosomes coated with Xist RNA. Therefore, DNA hypomethylation of the Xist gene may allow its ectopic expression with resultant transcriptional silencing of X-linked genes. This suggests that DNA methylation is not necessary for inactivation of X-linked genes. This agrees with other studies that indicated the DNA methylation occurs after X-linked genes are inactivated (89). Considering this and the evidence mentioned above, DNA methylation may be acting to maintain the differential expression of X-linked genes but is probably not involved in establishing inactivation.

Currently the mechanisms of action of DNA methylation are unclear. It has been known for some time that some transcription factors are unable to bind methylated DNA (34, 64) and that other transcription factors are unaffected by DNA methylation of their binding sites (5, 54, 57). This makes it unlikely that DNA methylation is acting solely by directly interfering with transcription factor binding. Two proteins that selectively bind methylated DNA, MeCP1 and MeCP2, (84, 100) have been identified. In vitro studies indicated that these proteins repress transcription (8). In these studies, if methylated promoters were transfected in the presence of competitive methylated DNA, expression was observed from the promoter. No expression was observed from methylated promoters which were transfected in the absence competitive methylated DNA. Transfection of methylated promoters into MeCP-1 deficient cells showed reduced repression of expression. Therefore, MeCPs may be binding to the methylated DNA of inactive X-linked genes and excluding the binding of transcriptional activators. However, in vivo footprinting analysis of X-linked promoters has failed to identify any factors bound to the inactive allele (63, 119, 122) which suggests that MeCPs are not required for stable repression. Another explanation is that these methods are unable to detect bound MeCPs. If DNA methylation is not acting directly to prevent transcription factor binding or indirectly by binding MeCPs, it may be controlling the differential accessibility of X-linked promoters via interaction with another mechanism such as chromatin structure. Though evidence for the involvement of DNA methylation in regulating the transcription of X-linked genes by X inactivation is strong, the mechanism by which DNA methylation acts and its precise role in X inactivation are unknown.

In order further to clarify the role of differential DNA methylation in the transcriptional silencing of X-linked genes, in this thesis, I have used genomic sequencing to examine the high-resolution DNA methylation pattern on the 5' region of the active and inactive mouse HPRT genes and compared these data to previous studies.

The Role of Chromatin Structure

Controlling the differential expression of genes on the active and inactive X chromosomes may also involve changes in chromatin structure. Comparison of the chromatin structure of the active and inactive X chromosomes reveals several differences. The presence of a Barr body in somatic cell nuclei of female cells suggests that the inactive X chromosome is maintained in a highly condensed heterochromatic state (112). Examination with DNase I indicates that chromatin structure of X-linked genes is different on the active and inactive X chromosome (52, 85, 126, 127, 153, 155). For the HPRT and PGK-1genes, the allele present on the active X chromosome is more sensitive to DNase I digestion than the inactive allele (127, 155), and the promoter regions of both genes contain a DNase I hypersensitive site only in the active promoter. Analysis of the entire X chromosome using in situ nick-translation with DNase I indicated that the active X chromosome is more susceptible to DNase I digestion that the inactive X chromosome (72). In vivo DNase I footprinting of the human PGK-1 and HPRT promoters (119) detected nucleosomes associated with the inactive promoter. On the active promoter

DNase I footprinting detected potential transcription factors but did not detect any nucleosomes. Furthermore there is a difference in the histone acetylation of the nucleosomes on the active and inactive X chromosome. Studies indicate that transcriptionally active chromatin contains acetylated histones and that transcriptionally silent chromatin contains non-acetylated histones (145). Examination of the histones acetylation of X-linked promoters indicates that the inactive promoters contain ordered unacetylated nucleosomes whereas the active promoters contain acetylated nucleosomes. Consequently, the process of X inactivation may involve regulating changes in the chromatin structure of genes on the active and inactive X chromosomes.

In order to clarify the role of chromatin structure and DNA methylation in the regulation of X-linked genes, in this thesis, I have examined the binding of transcription factors to the human HPRT gene, the nuclease-sensitivity, and the levels of HPRT mRNA at various times following the reactivation of the HPRT gene by treatment with the demethylating agent 5-azacytidine 2'dexoribonucleic acid (5aCdr).

X Chromosome Inactivation in Males

Unlike female mammals, male mammals have only one X chromosome. The male X chromosome is transcriptionally active in somatic tissues throughout development and into adulthood. However, it is thought that during spermatogenesis, at or just prior to the onset of meiosis (95, 97, 136), the male X chromosome undergoes a process of inactivation which may be similar to the process of X chromosome inactivation observed in females. Cytological evidence indicates the male X chromosome is heteropyknotic in meiotic spermatocytes (76, 111), it replicates late (76, 111), it is transcriptionally silent (74, 75, 108), and it has a hairpin-like accessory structure on the X axis (117) which is characteristic of a female inactive X chromosome. This evidence has been interpreted as X inactivation during male gametogenesis. Once inactivated, the sperm X chromosome remains transcriptionally silent in meiotic spermatocytes, postmeiotic spermatids and spermatoza, and does not become active until after fertilization (135). Cytological (76, 111), protein (77, 146), and RNA studies (74, 97, 136) suggest this inactivation occurs at, or just prior to prophase of the first meiotic division (75), thus significantly preceding the genome-wide silencing of genes associated with protamine replacement of histones and chromosome condensation during spermiogenesis in postmeiotic cells. However, the exact timing of inactivation during spermatogenesis in not currently known.

As observed in female X inactivation, the Xist gene may also to be playing a role the inactivation of the male X chromosome during spermatogenesis (96, 131). In males, Xist expression detected is only in the testis (96). No Xist transcript is detected in any other male tissue (96). Furthermore, Xist expression is undetectable until just prior to the onset of inactivation. This suggests that the male and female X chromosomes are inactivated by similar mechanisms. Like X inactivation in females, the molecular mechanisms of initiation, spreading, and maintenance of inactivation during spermatogenesis is not well understood. Therefore, understanding X chromosome inactivation in males may help clarify the mechanism involved in X inactivation in both males and females.

In order better to characterize the events of X inactivation in males, in this dissertation, I have used in vivo footprinting to determine the exact stage of spermatogenesis at which the mouse HPRT gene is inactivated or no longer bound by transcription factors.

HPRT Gene

It is probable that DNA-protein interactions, chromatin structure, and DNA methylation are all involved in the differential regulation of X-linked genes by X inactivation. Although X inactivation is a global chromosome-wide process, some measure of regulation is occurring at the level of individual genes. Therefore, understanding the molecular mechanism involved in the differential expression of individual genes on the active and inactive X chromosomes should give insight into whole process of X inactivation. Thus, in this dissertation, I examined the HPRT gene in order to gain a better understanding of the mechanisms of X inactivation.

Hypoxanthine phosphoribosyltransferase (HPRT) gene encodes a transcript that is constitutively expressed in all mammalian tissues with elevated expression in cells of the central nervous system (140). In the purine salvage pathway, HPRT converts hypoxanthine and guanine to inosine monophosphate (IMP) and gaunosine monophosphate (GMP), respectively (140). Because the mammalian HPRT gene is X-linked and subject to transcriptional silencing on the inactive X chromosome (46, 140), each female somatic cell nucleus carries both a transcriptionally active and inactive HPRT allele.

Similar to numerous constitutively expressed housekeeping genes (6, 11), the promoters of both the human and mouse HPRT genes are contained within a classic CpG island which contains multiple GC box sequences, lack TATA and CAAT sequences and employ multiple transcription initiation sites (101, 115). The human HPRT promoter contains multiple transcription start sites in a region from -104 to -169 (140). A region from -219 to -122 appears to be sufficient for normal expression levels of the HPRT gene

(140). Also, a negative regulatory element has been reported in the region from -570 to -388 (128). The mouse HPRT promoter contains one major transcription start site and two minor sites at +28 and +34 (101, 102). Normal expression of the mouse HPRT gene is reported to require only 49 bp of 5' flanking sequence (102). No evidence suggesting the presence of enhancer elements or negative regulatory elements has been reported in the mouse gene, though the human gene contains a negative regulatory element from position -485 to -474 of the 5' flanking region (33).

The Human XIST Gene

Since several lines of evidence indicate the XIST gene is involved in X inactivation, understanding the mechanisms which regulate the differential expression of the human XIST gene may provide insight into the mechanisms involved in the initiation of X inactivation. Therefore, in this dissertation, I have examined the differential binding of regulatory factors to the active and inactive XIST promoter using in vivo footprinting.

The XIST gene is the only gene to date that is expressed exclusively from the inactive X chromosome (14). In addition to its unique expression pattern, it has been mapped to the XIC (14). It expresses a ~17 kb RNA which contains no long, conserved open reading frame (15). The XIST gene contains two regions of repeats (15). The sequence between positions +6000 to +8300 contains a series ~300 bp direct repeats which show ~70% homology to the mouse Xist repeats. Between positions +350 to +770 a series of nine 43-59 bp repeats shows a strong sequence identity with the mouse Xist gene. The RNA does not appear to be translated into a protein product, nor does it appear to leave the nucleus. In situ hybridization studies indicate that the RNA transcript colocalizes to the inactive X chromosome in interphase cells (15). In vitro analysis of the human XIST promoter indicates that a region from -93 to +30 is necessary for high levels of XIST expression (56). Further in vitro mutagenesis studies coupled with transient expression assays and nucleotide sequence analysis indicate that the minimal promoter contains 5 regions important for expression, two potential Sp1 sites, a string of 7 cytosines, a TATA binding site (TBP), and a YY1 binding site.

Specific Aims and Goals

The goal of this dissertation is to examine the mechanisms of transcriptional regulation of X-linked genes as a result of mammalian X chromosome inactivation. Since X inactivation appears to involve multiple molecular mechanisms in the regulation of X-linked genes, the binding of sequence-specific DNA regulatory proteins, the DNA methylation pattern, and the timing of nuclease-sensitivity and transcription factor binding following 5aCdr-induced reactivation have been examined on the HPRT gene. Understanding the mechanisms that regulate the differential expression of an individual X-linked gene may provide insight into the mechanisms of the whole process of X inactivation. Furthermore, the timing of X inactivation of the male mouse HPRT gene during spermatogenesis has been determined using in vivo footprinting. Better characterization of the process of X inactivation in males may provide insight into the mechanisms involved in X inactivation for both males and females. Finally, the differential binding of regulatory factors to active and inactive human XIST alleles has been examined by in vivo footprinting. Because evidence suggests the XIST gene is involved in X inactivation, understanding the mechanisms regulating XIST expression may provide insight into the mechanisms of X inactivation.

The specific aims of this dissertation are briefly discussed below and presented in greater detail in the chapters. In chapter 2, the in vivo footprinting and high-resolution methylation analysis of the mouse HPRT gene 5' region on the active and inactive X chromosomes is presented. The 5' region of the mouse HPRT gene was examined by in vivo footprinting to identify potential DNA-binding regulatory proteins on the active and inactive alleles. These data have been compared with previous data from in vivo footprinting studies of the human HPRT gene. Genomic sequencing was used to determine the methylation state of each CpG dinucleotide in the 5' region of the mouse HPRT gene on the active and inactive X chromosomes. These data have been correlated with transcriptional activity and with differences observed in the in vivo footprinting studies. In addition, the high-resolution methylation patterns of the mouse and human HPRT gene have been compared. In chapter 3, the in vitro reconstitution of a DNA-protein interaction that occurs over a novel nine base pair sequence which is identical in both the mouse and human HPRT genes is presented. Gel mobility-shift assays were performed using crude Hela nuclear extract and double stranded oligonucleotides that spanned the novel nine base pair sequence. In chapter 4, nuclease-sensitivity, transcription factor binding, and mRNA have been examined for the human HPRT gene following 5aCdr treatment. These experiments attempt to investigate the relationship between chromatin structure, DNA methylation, and transcription factor binding following reactivation of an inactive X-linked genes. In chapter 5, in vivo footprinting of the mouse HPRT gene has been examined at different stages of spermatogenesis. These experiments follow the binding of transcription factors identified in chapter 2 to determine the stage during which the male X chromosome becomes inactivated. In chapter 6, in vivo footprinting of the active and inactive human XIST gene is presented. These studies examined the differential binding of regulatory factors to the active and inactive human XIST gene in order to identify potential DNA-binding regulatory proteins. In chapter 7, the conclusions obtained from these studies and future experimental directions are presented.

CHAPTER 2

IN VIVO FOOTPRINTING AND HIGH RESOLUTION METHYLATION ANALYSIS OF THE MOUSE HYPOXANTHINE PHOSPHORIBOSYLTRANSFERASE GENE 5' REGION ON THE ACTIVE AND INACTIVE X CHROMOSOMES

Introduction

Hypoxanthine phosphoribosyltransferase (HPRT; EC 2.4.2.8) operates in the purine salvage pathway by converting hypoxanthine and guanine to their respective nucleotides, IMP and GMP (140). It is present in all mammalian tissues at all stages of development and its expression is elevated in cells of the central nervous system (140). The mammalian HPRT gene is X-linked and subject to transcriptional silencing on the inactive X chromosome in female somatic cells by X chromosome inactivation (46, 140). Thus, each female somatic cell nucleus carries both a transcriptionally active and inactive HPRT allele.

The promoters of both the human and mouse HPRT genes lack TATA and CAAT sequences and employ multiple transcription initiation sites (101, 115). In addition, both promoters are contained within a classic CpG island and contain multiple GC box sequences, features common to numerous constitutively expressed housekeeping genes, both autosomal and X-linked (6, 11). Normal expression of the mouse HPRT gene is reported to require only 49 bp of 5' flanking sequence (102). No evidence suggesting the presence of enhancer elements or negative regulatory elements has been reported in the mouse gene, though the human gene contains a cell-type-specific negative regulatory element from position -485 to -474 of the 5' flanking region (128).

The transcriptional inactivation of genes on one of the two X chromosomes in female somatic cells by X chromosome inactivation equalizes the dosage of functional X-linked genes between male and female somatic cells (46, 48). The molecular mechanisms responsible for initiating, spreading, and maintaining X inactivation at the chromosomal level are unknown, as are the mechanisms regulating inactivation of individual X-linked genes by this process. DNA-protein interactions (46, 94), chromatin structure (72, 112, 119) DNA replication (45, 46), and DNA methylation (46, 86, 107) have all been postulated to be involved. Though evidence is strong for the involvement of DNA methylation in regulating transcription of X-linked genes by X inactivation, the mechanism by which DNA methylation acts and its precise role in X inactivation are unknown. Transcriptional regulation at the level of individual X-linked genes (or chromosomal domains) by DNA methylation is suggested by experiments using the demethylating agent 5-azacytidine (5-azaC) which has been shown to independently reactivate individual genes (including the HPRT gene) on the inactive X chromosome (53, 107, 148). Thus, analysis of the transcriptional regulation of individual X-linked genes by X inactivation is likely to provide insight into the molecular mechanism(s) of this process. Because the HPRT gene has been extensively characterized in both mouse and man and is subject to X inactivation in both species, a comparison of conserved human and murine regulatory motifs associated with the active and inactive alleles may help identify mechanisms by which individual X-linked genes are differentially expressed on the active and inactive X chromosomes.

Previously, we examined the binding of transcription factors and the high resolution DNA methylation pattern of the human HPRT promoter region by in vivo footprinting and genomic sequencing, respectively, on the active and inactive X chromosomes (61, 63). In vivo footprinting with dimethyl sulfate (DMS) indicates the 5' flanking region of the active allele was occupied by multiple DNA-binding transcription factors, while the inactive allele exhibited no detectable sequence-specific DNA-binding proteins. The in vivo footprints on the active allele include those associated with five potential SP1 binding sites, a potential AP-2 binding site, and a sequence just downstream of the transcription initiation region. These studies, in concert with similar studies of the the X-linked human phosphoglycerate kinase (PGK-1) gene (119, 122) argued against models for X inactivation that require the existence of promoter elements unique to X-linked genes that bind trans-acting activator and/or repressor proteins and coordinately regulate transcription of genes on the active and inactive X chromosomes (63, 94).

High resolution DNA methylation analysis of the promoter region of the human HPRT gene by genomic sequencing showed that every cytosine in the 5' region of the active allele was unmethylated. The same region of the inactive allele was completely methylated at nearly every CpG dinucleotide, except within a 68 bp region associated with the five potential Sp1 binding sites which contains numerous CpG dinucleotides that are unmethylated or partially methylated (61). The basis for this unusual methylation pattern of the inactive allele is unknown, but may be related to the recent report that Sp1 binding sites protect CpG islands from de novo methylation (10).

Until now, the differential binding of regulatory factors to the active and inactive alleles of an X-linked gene in a non-human mammal has not been examined. Therefore, the in vivo binding pattern of transcription factors observed for the human HPRT and PGK-1 genes on the active and inactive X chromosomes has not been demonstrated to be conserved in other mammalian species. In addition, examination and comparison of the high resolution DNA methylation pattern of X-linked genes in various mammalian species may reveal conserved methylation patterns critical for the differential transcription of genes by X chromosome inactivation.

To examine the conservation of transcription factor binding patterns and high resolution methylation patterns between the human and mouse HPRT gene 5' regions, we have now performed ligation-mediated PCR (LMPCR)-assisted in vivo footprinting and genomic sequencing analysis of the mouse HPRT gene on the active and inactive chromosomes. We employed a novel method in conjunction with LMPCR to separate the active and inactive HPRT alleles in female somatic cells for in vivo footprinting and genomic sequencing. We found a high degree of conservation in the interaction of transcription factors with the human and mouse genes. Furthermore, we identified a previously unreported 9 bp sequence that was perfectly conserved and identically footprinted in both species that may act as a transcriptional initiator element. However, the unusual pattern of DNA hypomethylation associated with a cluster of GC boxes in the human HPRT gene on the inactive X chromosome is not conserved in the inactive mouse gene (though the overall feature of hypermethylation of the 5' region is retained on the inactive allele in both species).

Materials and Methods

Cell Lines

Three cell lines, B5, C3 and 3T3-L1, were used for in vivo footprinting and genomic sequencing. B5 and C3 cells were prepared from embryos of an interspecies hybrid between *Mus musculus* and *Mus caroli* by Chapman and Shows (30). The B5 cell line contains an active *M. musculus* X chromosome and an inactive *M. caroli* X chromosome, and the C3 cell line contains an inactive *M. musculus* X chromosome and an active *M. caroli* X chromosome as determined by isozyme analysis (30). 3T3-L1 cells were derived from a single-cell clone of mouse 3T3 fibroblasts and contain an active X chromosome (50).

Cells were grown in Dulbecco's modified eagle medium (D-MEM) with 10% fetal bovine serum (FBS) and 1% Penicillin-streptomycin (P-S) to confluence in T-150 flasks.

Dimethyl Sulfate (DMS) Treatment

In vivo DMS treatment of the three cell lines was performed as previously described (63). For in vitro DMS treatment, genomic Taq I DNA fragments containing the 5' region of the active and inactive *M. musculus* HPRT alleles were isolated from B5 and C3 cells (see method below), and 50 ug of the purified DNA were treated with 0.5% DMS for 1 minute at room temperature. Prior to in vitro DMS treatment of total 3T3-L1 genomic DNA, the purified genomic DNA was digested with Hind III to reduce viscosity.

Separation and Isolation of the Active and Inactive 5' Regions of the *M. musculus* HPRT Gene

The 5' regions of the HPRT genes from *M. musculus* and *M. caroli* in B5 and C3 hybrid were separated by taking advantage of a species-specific Taq I restriction fragment length polymorphism (RFLP) identified by Locke et al. (89). The 5' region of the *M. musculus* HPRT gene is contained in a 5.3 kb Taq I genomic DNA fragment, whereas the 5' region of the *M. caroli* HPRT gene is contained within a 3.0 kb Taq I fragment. Therefore, isolation of genomic DNA fragments approximately 5.3 kb in size from Taq I-digested genomic DNA from B5 cells selects for the 5' region of the HPRT gene from the active *M. musculus* X chromosome; isolation of 5.3 kb fragments from Taq I digested genomic DNA from C3 cells selects for the 5' region of the HPRT gene from the inactive *M. musculus* X chromosome (89, 155).

To separate physically the 5' region of the *M. caroli* and *M. musculus* HPRT alleles, 60 ug of purified genomic DNA from DMS-treated and untreated B5 and C3 cells were digested with 1200 units of Taq I for 8 hrs at 37°C. DNA was digested at 37°C instead of the optimum temperature of 65°C for Taq I to prevent depurination of DMS-modified guanines. Taq I-digested genomic DNA was then size-fractionated by electrophoresis on a 1% preparative agarose gel. Using an adjacent lane of DNA size markers for guidance, a slice of agarose containing DNA fragments from 4 to 9 kb of the Taq I-generated genomic DNA fragments was excised from the gel. This slice included the 5.3 kb Taq I fragment containing the 5' region of the *M. musculus* HPRT gene and lacked the 3.0 kb *M. caroli*-specific fragment. DNA was eluted from the agarose slices by electroelution into dialysis bags as described by Sambrooke et al. (132) using autoclaved 0.5X TBE (50 mM Tris-HCl, 43mM boric acid, 0.5 mM EDTA, pH 8.3) for both the gel buffer and the elution buffer. Eluted DNA was concentrated with butanol, extracted once with phenol, once with phenol:chloroform:isoamyl alcohol (PCI 25:24:1), and once with chloroform. After extraction, DNA was precipitated with 2 volumes of ethanol and collected by centrifugation.

Following isolation of the size-selected Taq I fragments containing either the active (from B5 cells) or inactive (from C3 cells) *M. musculus* 5' region, the DMS-modified DNA was cleaved by treatment with 100 ul of 10% piperidine for 30 minutes at 95°C as described (63). Control naked genomic DNA isolated from cells not treated with DMS was modified in vitro by DMS treatment and then subjected to piperidine cleavage as described above. After piperidine cleavage, the DNA was dried overnight in a vacuum concentrator to remove piperidine and resuspended in 20 ul TE (10 mM Tris-HCl, pH-8, 1 mM EDTA).

Hydrazine Treatment for Genomic Sequencing

Fifty ug of purified DNA from 3T3-L1 cells, or genomic DNA containing the active or inactive 5' region of the *M. musculus* HPRT gene (prepared as described above) were resuspended in 5 ul water + 15 ul 5 M NaCl. The DNA was then subjected to the standard Maxam and Gilbert cytosine-specific modification/cleavage reaction with hydrazine and piperidine (93). Hydrazine treatment for 10 minutes was found to be optimal. Following cleavage of the hydrazine-modified DNA with piperidine, 1 ml of water was added and the sample was dried overnight in a vacuum concentrator. The chemically cleaved DNA was then resuspended in TE to a final concentration of 1 ug/ul.

Preparation of the Control DNA Sequencing Ladder

To generate a complete sequencing ladder of the 5' region of the mouse HPRT gene, the plasmid pHPT6 (101) containing a 1.4 kb EcoRI fragment of the mouse HPRT 5' region in pBR322 was first digested with EcoRI. Two and one half ug of digested plasmid DNA was then chemically modified and cleaved by standard G-, G+A-, T+C-, C-specific Maxam-Gilbert DNA sequencing reactions (93). Cleaved plasmid DNA was diluted (approximately 500 fold) to a concentration which produced the same signal intensity on an autoradiograph as PCR-amplified genomic DNA samples.

Ligation-Mediated PCR

The *M. musculus* HPRT 5' region in all chemically modified and cleaved DNA (for both in vivo footprinting and genomic sequencing) was amplified by ligation-mediated PCR (LMPCR) (63). For LMPCR, seven oligonucleotide primer sets were synthesized according to the published mouse HPRT sequence (101). For analysis of the lower strand, the primer sets were as follows: set R primers 1 and 2, CTAGCCAGACTCCATGAATC and AATCCCTGCCCGCCCCGCCATCAAA, respectively; set S primers 1, 2 and 2a, TGTGGGGCTCTGCTGGAGT, TGGAGTCCCCTTGGCTCACCACGA, and GCGGCAAAAAGCGGTCTGAGGAGGAA, respectively; set T primers 1 and 2, GCTCCGCTGGGCGTTCT and CCACCAAAGGCAGTTCCGGAACTC, respectively; and set U primers 1 and 2, TCCGGAAAGCAGTGAGGTAA and TGAGGTAAGCCAACGCTCTCCCTC, respectively. For analysis of the upper strand, the primer sets were as follows: set G primers 1 and 2, ATTAAAGGCGTGCGCCACCA and CACCGCCCGGCTGGATCTCAAATCT, respectively; set H primers 1 and 2, TGCGGAGTGATTATCTGGGA and CTGGGAATCCTCTGGGAGACGACA, respectively; set I primers 1 and 2, AGCGTTTCTGAGCCATTGCT and GAGGGAGAGCGTTGGGCTTACCTCA, respectively; and set J primers 1 and 2, TGCGGAGTGATTATCTGGGA (same as primer 1 of primer set H) and GAGTTCCGGAACTGCCTTTGGTGGC, respectively. Figure 2.1 shows the relative position and region covered by each of the primer sets for the 5' region of the mouse HPRT gene.

LMPCR was performed as described in Yang and Hornstra (61). For PCR, DNA was denatured at 95°C for 20 seconds, annealed at 64°C for 1 minute, and extended at 76°C for 1 minute with an additional 5 seconds each cycle for 20 cycles (in a Coy Temperature Cycler II). LMPCR in some in vivo footprinting studies was carried out using Taq polymerase (rather than VENT polymerase) during the amplification step according to Hornstra and Yang (63).

Gel Electrophoresis and Electrotransfer

Visualization of all sequencing ladders was performed by the electrotransfer and hybridization method described by Hornstra and Yang (63). Gel electrophoresis and transfer were carried out as previously described (62, 63).

Probe Synthesis, Hybridization, and Washing

^{32}P-labelled strand-specific hybridization probes were used to visualize the DNA sequencing ladders as previously described (62, 63). The strand-specific hybridization probes were synthesized from a single-stranded M13 template by first cloning the 1.4 kb EcoRI mouse HPRT 5' genomic DNA fragment from plasmid pHPT6 (101) into the EcoRI site of the M13 cloning vector M13mp18. This yielded two subclones with the

HPRT insert in different orientations (verified by DNA sequencing) and allowed the isolation of the single-stranded templates used for synthesis of probes specific to the upper and lower strands of the 5' region of the mouse HPRT gene. Large-scale preparations of each single-stranded template in M13 were generated as described by Sambrook et al. (132). Probe synthesis, hybridization and washing were performed as described previously (62, 63).

Results

To examine mechanisms which regulate the differential expression of genes by X chromosome inactivation, LMPCR-assisted in vivo footprinting and genomic sequencing were used to analyze the 5' region of the mouse HPRT gene on the active and inactive X chromosomes. The regions from positions -305 to +205 of the lower stand and positions -279 to +234 of the upper strand were examined by in vivo footprinting to identify sequence-specific DNA-protein interactions unique to either the active or inactive HPRT alleles. In addition, the region from positions -270 to +267 of the upper strand and positions -207 to +284 of the lower strand were subjected to analysis by genomic sequencing to determine the methylation state of every cytosine nucleotide on the active and inactive alleles. These studies covered approximately 500 bp of the 5' region and included the minimal promoter of the mouse gene as determined by Melton et al. (102), 190 bases 5' of the minimal promoter, the 5' untranslated region, the first exon, and the 5' end of the first intron. These studies allowed an examination of correlations between the high-resolution methylation pattern and in vivo footprinting pattern of the mouse promoter region, as well as comparisons with results from similar studies of the human

HPRT promoter (61, 63), and the X-linked human and mouse phosphoglycerate kinase (PGK-1) gene promoters (122, 143).

In Vivo Footprinting

The relative locations of the oligonucleotide primer sets used for LMPCR in vivo footprinting of the 5' region of the mouse HPRT gene are shown in Figure 2.1. Primers were designed and positioned such that data from one primer set would overlap with data from the preceding primer set and were specific for the *M. musculus* DNA sequence.

Three mouse cell lines were analyzed, B5, C3, and 3T3-L1. The B5 and C3 cell lines are derived from an interspecies hybrid cross between *M. musculus* and *M. caroli.* Isoenzyme analysis indicates that the *M. musculus* X chromosome is active in the B5 cell line and inactive in the C3 cell line (30). To separate the active and inactive HPRT alleles in C3 and B5 cells, we have employed a previously reported (89) species-specifc Taq I polymorphism that permits discrimination of the 5' regions of the *M. musculus* and *M. caroli* HPRT genes; the *M. musculus* 5' region is contained within a 5.3 kb Taq I fragment and the *M. caroli* 5' region is contained in a 3.0 kb Taq I fragment. Thus, isolation of Taq I genomic DNA fragments approximately 5.3 kb in size will selectively enrich for the active *M. musculus* allele in B5 cells and the inactive *M. musculus* allele in C3 cells (see Materials and Methods). Footprints detected in B5 cells from DMS-treated DNA isolated in this manner should be those associated with the active allele, while footprints detected in size-selected DNA from the C3 cell line should represent those from the inactive allele. The *M. caroli* gene was not analyzed because the nucleotide sequence of the 5' region from this species has not been determined.

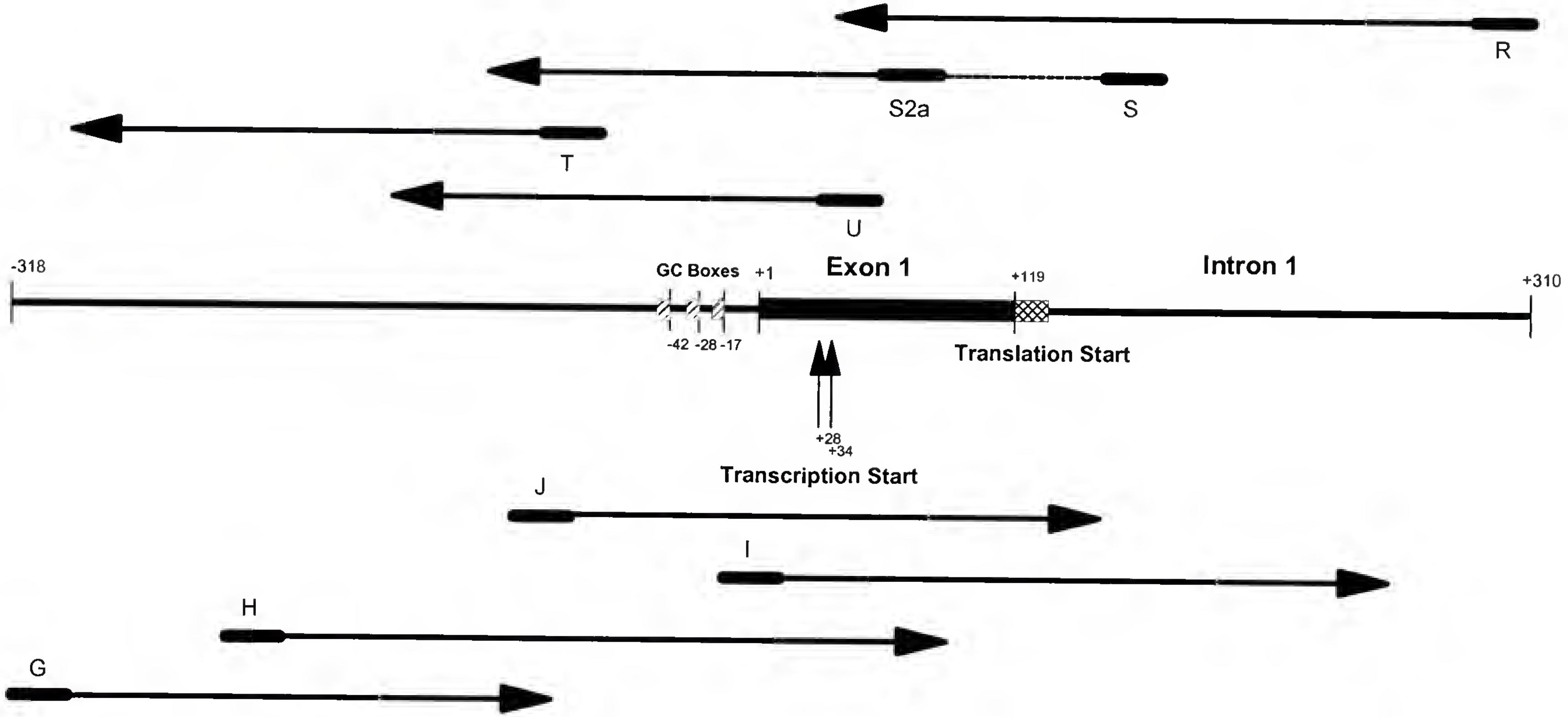
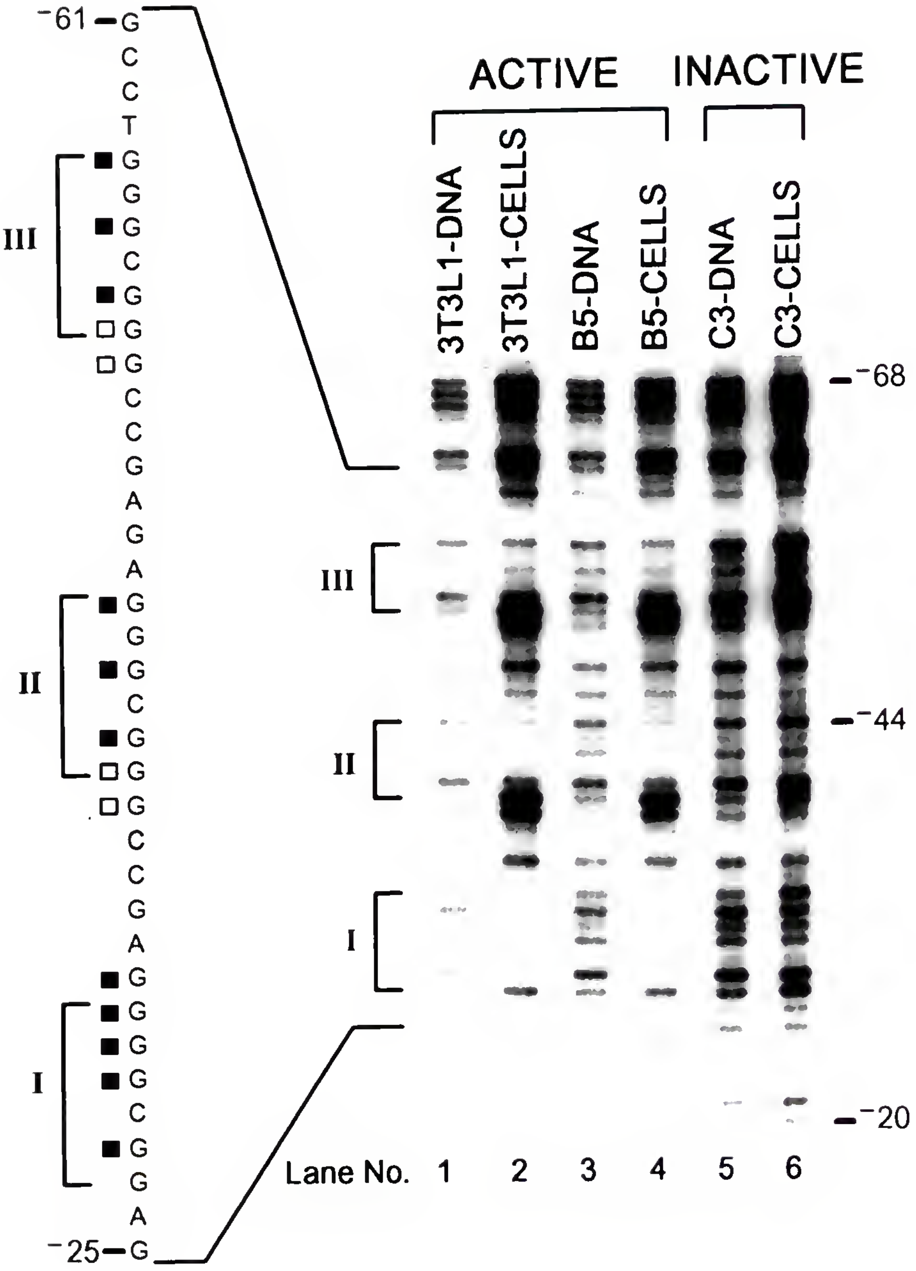


FIGURE 2.1 - Location of LMPCR primers for analysis of the mouse HPRT 5' region. The numbered center line represents the mouse HPRT gene 5' region with positions numbered relative to the primary transcription initiation site. The large filled rectangle that begins at position +1 represents the first exon with the solid portion signifying the 5' untranslated region and the stippled area signifying the translated region. The smaller diagonally-striped boxes represent GC boxes (GGGCGG). The large horizontal arrows above and below the numbered line indicate the positions of the PCR primer sets used for LMPCR and the approximate region and direction examined by each primer set. Primer sets G, H, I, and J were used to analyze the lower strand. Primers sets R, S, T, and U were used to analyze the upper strand. S2a is an additional primer used to expand the region covered by the S primer set. The dotted line represents the region covered between primer S and S2a. The vertical arrows indicate positions of the minor transcription initiation sites.

Because the B5 and C3 cell lines are likely to carry an autosomally-encoded component of *M. caroli* transcription factors, the 3T3-L1 mouse cell line was analyzed as a control for *M. musculus* transcription factors interacting with an active *M. musculus* promoter. This cell line was subjected to selection in HAT medium prior to in vivo footprint analysis and therefore contains an active mouse X chromosome; consequently, in vivo footprints observed in the 3T3-L1 cell line should be representative of DNA-protein interactions on the active mouse HPRT allele.

Representative autoradiograms from analysis of the *M. musculus* 5' region in these three cell lines are shown in Figures 2.2 and 2.3. Analysis of the upper strand of the mouse HPRT gene with primer set U is shown in Figure 2.2. This primer set covers the primary transcription initiation site and a series of three GC boxes with sequences analogous to the human Sp1 binding sequence (11). Comparison of the DMS modification and cleavage pattern between naked DNA and in vivo treated DNA revealed footprinted regions in the 3T3-L1 and B5 samples treated in vivo with DMS (comparing lane 1 with lane 2, and lane 3 with lane 4). These footprints occur over the three GC boxes which are indicated by the brackets in Figure 2. For GC box I, all footprinted guanine nucleotides in the in vivo treated samples (lanes 2,4) are protected from reaction with DMS relative to the DMS modification and cleavage pattern of naked DNA (lanes 1,3). In GC box II and III, the last two residues exhibit enhanced DMS reactivity and the other footprinted guanine nucleotides are protected from reaction with DMS. Footprints observed at guanine residues -57, -55, -53, -51, -44, -42,-40 (relative to the +1 major transcription start site) were verified in other autoradiograms using primer set S (data not shown). Since the size-selected DNA samples from 3T3-L1 and B5 cells in Figure 2 both

FIGURE 2.2 - In vivo footprint analysis of the upper strand using primer set U. Autoradiogram of the guanine-specific sequencing ladder from DMS treatment. The nucleotide sequence of the region is shown to the left of the guanine-specific sequencing ladder generated by DMS; the position of each nucleotide is indicated relative to the major transcription initiation site. Numbers to the right of the autoradiogram indicate positions of bands relative to the major transcription initiation site. Open squares to the left of the nucleotide sequence represent sites of enhanced DMS reactivity, and solid squares represent guanine nucleotides protected from reaction with DMS. Brackets with Roman numerals indicate the positions of the three GC boxes. Lane 1, naked DNA purified from 3T3-L1 cells and treated in vitro with DMS; lane 2, DNA from intact 3T3-L1 cells treated in vivo with DMS; lane 3, naked DNA purified from B5 cells and treated in vitro with DMS; lane 4, DNA from intact B5 cells treated in vivo with DMS; lane 5, naked DNA purified from C3 cells and treated in vitro with DMS; lane 6, DNA from intact C3 cells treated in vivo with DMS. Active indicates samples containing an active *M. musculus* HPRT promoter; inactive indicates samples containing an inactive *M. musculus* HPRT promoter. Cell lines are described in Materials and Methods. Guanine residue -43 in GC box II and residue -56 in GC box III were not resolvable due to compression. Guanine residue -40 in GC box II and residues -55 and -57 in GC box III were shown to be protected from DMS in other autoradiograms not shown.


⁻61
G
C
C
T
■ G
G
■ G
C
■ G
□ G
□ G
C
C
G
A
G
A
■ G
G
■ G
C
■ G
□ G
□ G
C
C
G
A
■ G
■ G
■ G
■ G
C
■ G
G
A
⁻25
G
III
II
I
ACTIVE
INACTIVE
3T3L1-DNA
3T3L1-CELLS
B5-DNA
B5-CELLS
C3-DNA
C3-CELLS
III
II
I
⁻68
⁻44
⁻20
Lane No.
1
2
3
4
5
6

contain the active mouse HPRT allele, these footprints indicate DNA-binding regulatory proteins interact with the active mouse HPRT promoter, an observation that is in accord with previous findings from the human HPRT gene (63).

In C3 cells (lanes 5 and 6), there is no detectable and consistent difference in the modification and cleavage pattern between the in vitro and in vivo DMS-treated samples (lanes 5 and 6, respectively). Because the size-selected samples from C3 cells contain only the inactive HPRT allele, the lack of footprints in this sample suggest that no sequence specific DNA-binding proteins are interacting within the 5' region of the inactive mouse HPRT gene, a result also consistent with studies of the human HPRT gene. The DMS modification and cleavage pattern of naked genomic DNA from all three cell types (Fig. 2.2, lanes 1, 3, 5) was identical to that of the same region cloned into a plasmid where purified plasmid DNA was subjected to the standard DMS and piperidine treatment followed by LMPCR (data not shown).

Examination of the lower strand from position -71 to +115 with primer set J (Figure 2.3) reveals two footprints, at positions -24 and +14. As seen on the upper strand (Fig. 2.2), these footprints were only associated with samples treated in vivo with DMS and containing the active 5' region (3T3-L1 and B5 cells, lanes 2 and 4). The protected guanine at -24 is located within GC box I. Neither footprint is detectable in the sample from C3 cells where the inactive allele is assayed (lane 6).

The DMS in vivo footprint pattern for the mouse HPRT 5' region on the active HPRT allele is summarized in Figure 2.4. Primer sets G, H, I, R, S, and T did not detect additional footprints on either the active or inactive alleles (data not shown). Sequence-specific DNA-protein interactions were observed only in association with the three GC

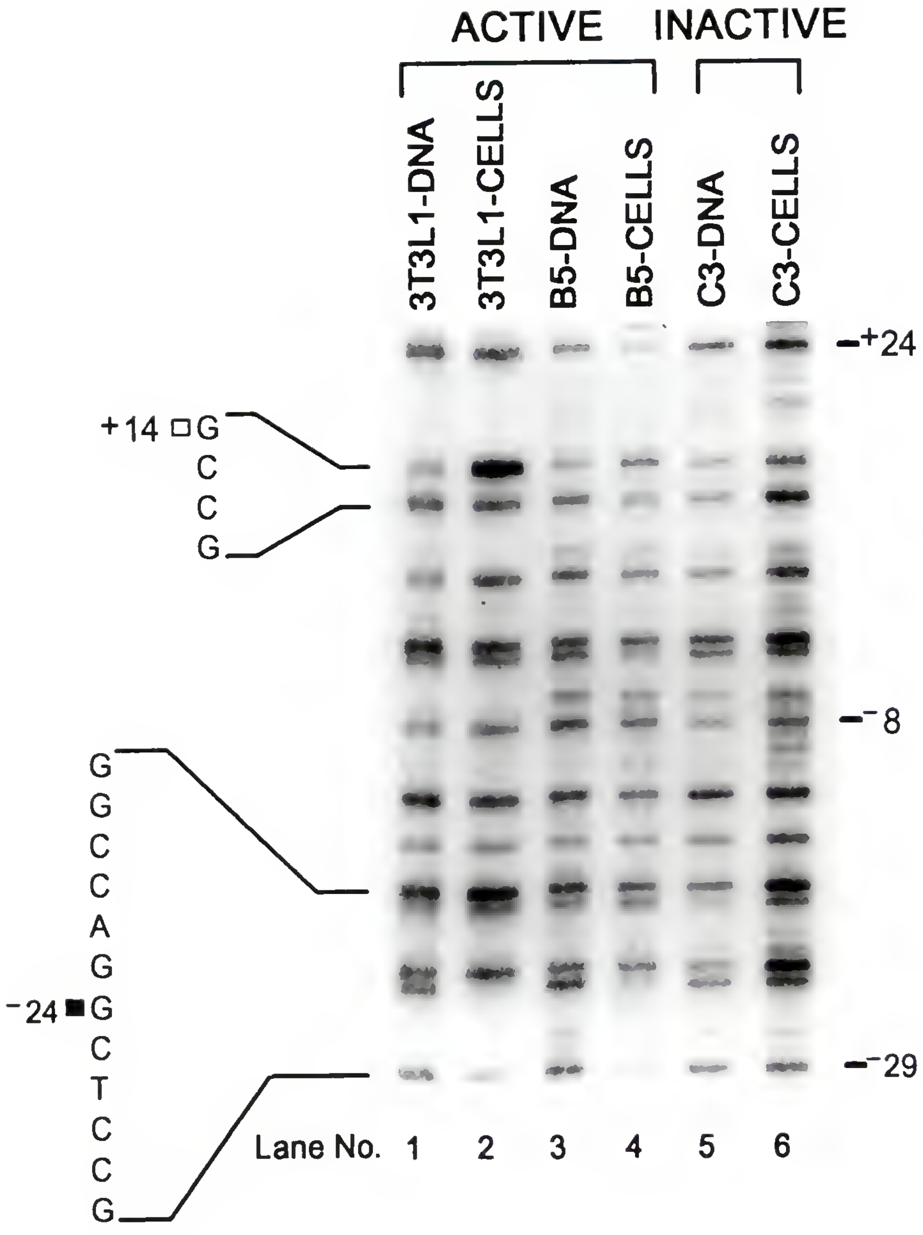


FIGURE 2.3 - In vivo footprint analysis of the lower strand using primer set J. All symbols and designations are identical to those described in Fig. 2.2.

boxes and just downstream of the transcription initiation site at position +14 (Figs. 2.2 and 2.3). Furthermore, these footprints on the active *M. musculus* gene in 3T3-L1 and B5 cells were identical, indicating that DNA-binding proteins from *M. caroli* are likely to behave identically to their *M. musculus* homologues in the C3 and B5 hybrid cells. As previously seen with the human HPRT and PGK-1 genes (63, 122), in vivo footprints were observed only on the active alleles (in B5 and 3R3-L1 cells). No footprints were detected on the inactive allele (in C3 cells), indicating that transcription factor binding occurs only on the active allele and that these factors are most likely to be transcriptional activators.

Comparison of the in vivo footprint patterns from the active allele of the mouse and human HPRT genes shows that the pattern is highly conserved in the two species. The transcriptionally active alleles of both genes carry footprints associated with a cluster of multiple Sp1 binding sites (5 sites in the human gene and 3 sites in the mouse) and with a region closely associated with the transcription initiation site(s). Neither species displays a detectable footprint on the inactive HPRT allele. The one notable difference between the mouse and human genes is that the active human gene exhibits an additional in vivo footprint upstream of the clustered GC boxes in a potential AP2 binding site (63), a DNA sequence (and footprint) that is not present in this region of the mouse gene.

Of particular interest in comparing the human and mouse in vivo footprint patterns is the unusual footprint near the transcription start site that is present in both active genes. This footprint in both species is detected only as sites of enhanced DMS reactivity, at position +14 (relative to the major transcription start site) in the mouse gene

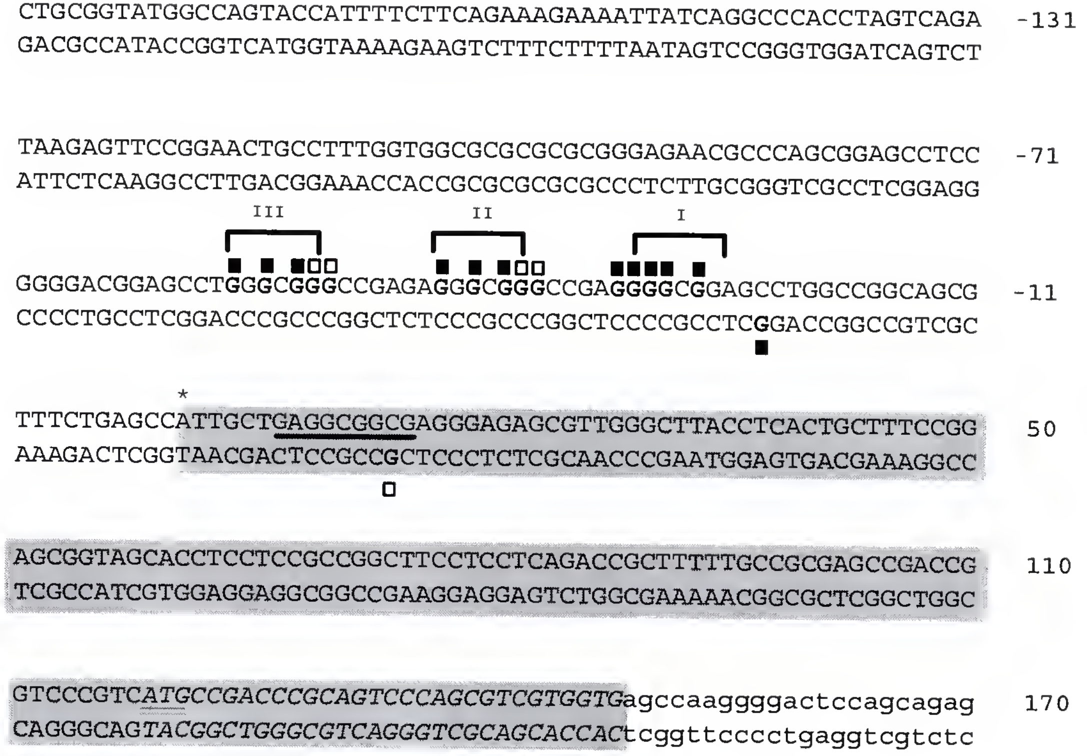
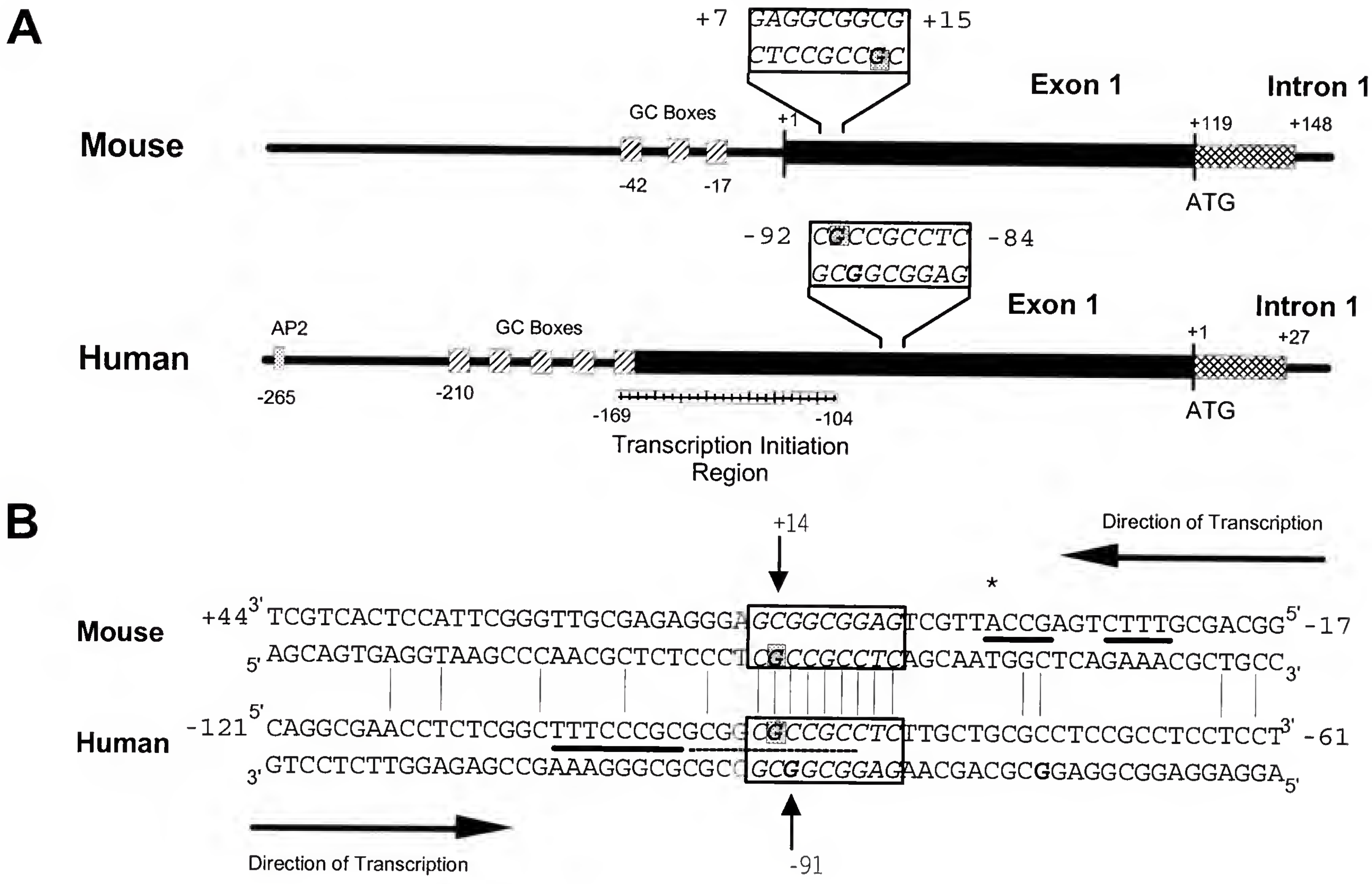

CTGCGGTATGGCCAGTACCATTTTCTTCAGAAAGAAAATTATCAGGCCCACCTAGTCAGA -131
GACGCCATACCGGTCATGGTAAAAGAAGTCTTTCTTTTAATAGTCCGGGTGGATCAGTCT
TAAGAGTTCCGGAACTGCCTTTGGTGGCGCGCGCGGGAGAACGCCCAGCGGAGCCTCC -71
ATTCTCAAGGCCTTGACGGAAACCACCGCGCGCGCCCTCTTGCGGTCGCCTCGGAGG
III
II
I
GGGGACGGAGCCTGGGCGGGCCGAGAGGGCGGGCCGAGGGGCGGAGCCTGGCCGGCAGCG -11
CCCCTGCCTCGGACCCGCCCGGCTCTCCCGCCCGGCTCCCCGCCTCGGACCGGCCGTCGC
*
TTTCTGAGCCATTGCTGAGGCGGCGAGGGAGAGCGTTGGGCTTACCTCACTGCTTTCCGG 50
AAAGACTCGGTAACGACTCCGCCGCTCCCTCTCGCAACCCGAATGGAGTGACGAAAGGCC
AGCGGTAGCACCTCCTCCGCCGGCTTCCTCCTCAGACCGCTTTTTGCCGCGAGCCGACCG 110
TCGCCATCGTGGAGGAGGCGGCCGAAGGAGGAGTCTGGCGAAAAACGGCGCTCGGCTGGC
GTCCCGTCATGCCGACCCGCAGTCCCAGCGTCGTGGTGagccaaggggactccagcagag 170
CAGGGCAGTACGGCTGGGCGTCAGGGTCGCAGCACCACtcggttcccctgaggtcgtctc

FIGURE 2.4 - Summary of in vivo footprints in the 5' region of the active M. musculus HPRT gene. Guanine nucleotides shown in bold exhibit in vivo footprints. Solid squares denote guanine nucleotides protected from DMS modification, and open squares denote guanine nucleotides exhibiting enhanced DMS reactivity. The shaded region indicates the first exon with italicized letters denoting the translated coding portion; the translation start site is indicated with a thin double line. Lower case letters denote the beginning of first intron. The three GC boxes are indicated by brackets with Roman numerals (corresponding to the Roman numerals and brackets in Figs. 2.2 and 2.3). Numbering is relative to the major transcription start site at position +1 denoted by the asterisk. The bold underline beginning at position +7 indicates the position of the conserved footprinted 9 bp sequence that is similarly footprinted in the human HPRT gene (see Fig. 2.5 and text).

FIGURE 2.5 - Alignment of the 5' region from the mouse and human HPRT genes. A. Diagrams of the 5' region of the mouse and human HPRT genes. The mouse HPRT gene and is numbered relative to the major transcription start site which is designated by +1. The human HPRT gene is numbered relative to the translation start site which is designated by a +1. The thick horizontal bar represents the first exon, with the solid portion indicating the 5' untranslated region and the cross-hatched portion indicating the translated portion. The region of multiple transcription initiation sites in the human HPRT gene is designated by a checkered bar below the 5' region of the first exon. The small hatched boxes show the positions of GC boxes, and the small shaded box on the human gene indicates the putative AP2 site. The 9 bp nucleotide sequence which is identical in the human and mouse promoters is shown within the large boxes and its position in each gene is indicated by brackets. The boxed and bolded G within the 9 bp sequence denotes the nucleotide which is footprinted by DMS in both the human and mouse HPRT genes. B. Nucleotide sequence alignment of the region containing the conserved 9 bp sequence. *Mouse* indicates the double-strand nucleotide sequence of the *M. musculus* HPRT promoter and is numbered relative to the major transcription initiation site. *Human* designates the double-strand nucleotide sequence of the human HPRT promoter and is numbered relative to the translation start site. The two sequences are aligned in reverse orientation relative to each other; the large arrows indicate the direction of transcription for each gene. The asterisk indicates the mouse major transcription initiation site (position +1). Thin vertical lines denote nucleotide sequence identity between the mouse and human genes in this alignment. The italicized and boxed nucleotide sequence denotes the 9 bp sequence which is identical (in anti-parallel orientation) in both the human and mouse HPRT promoter and similarly in vivo footprinted. The consensus binding sequence for the transcription factor HIP1 is underlined. Bolded and shaded G's denotes the guanine which is footprinted in both the human (position -91) and mouse (position +14) genes in the same location within the 9 bp sequence. Nucleotides shown in bold denote the guanines which exhibit enhanced DMS reactivity. The consensus binding sequence for the transcription factor HIP1 is underlined with solid lines. A palindromic sequence in the human gene is underlined with a dashed line.

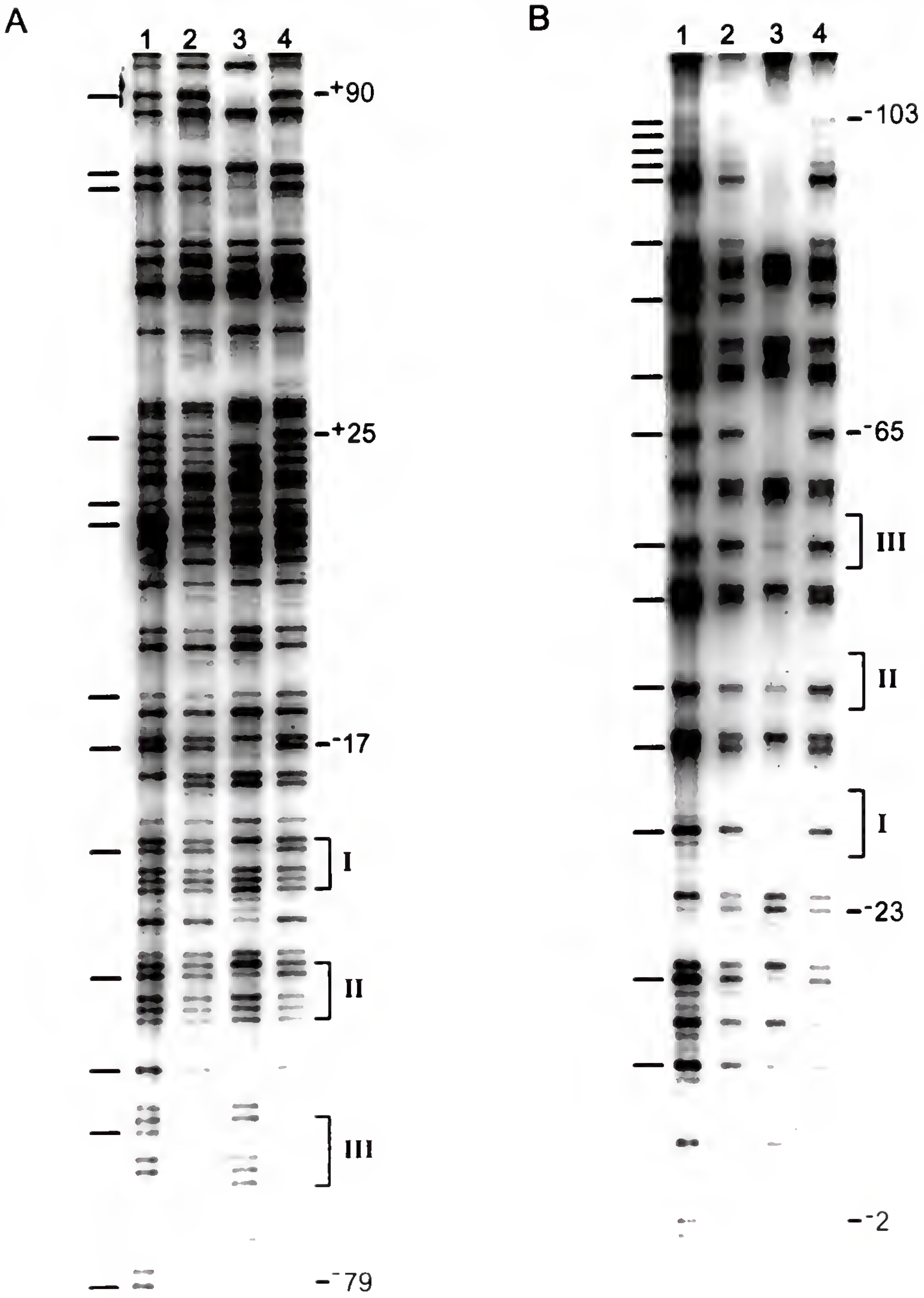

A
Mouse
+7
GAGGCGGCG
CTCCGCCGC
+15
Exon 1
Intron 1
GC Boxes
+1
+119
+148
-42
-17
ATG
Human
-92
CGCCGCCTC
GCGGCGGAG
-84
Exon 1
Intron 1
AP2
GC Boxes
+1
+27
-265
-210
-169
-104
ATG
Transcription Initiation Region
B
+14
Direction of Transcription
*
Mouse
+44 3' TCGTCACTCCATTCGGTTGCGAGAGGGAGCGGCGGAGTCGTTACCGAGTCTTTGCGACGG 5' -17
5' AGCAGTGAGGTAAGCCCAACGCTCTCCCTCGCCGCCTCAGCAATGGCTCAGAAACGCTGCC 3'
Human
-121 5' CAGGCGAACCTCTCGGCTTTCCCGCGCGGCGCCGCCTCTTGCTGCGCCTCCGCCTCCTCCT 3' -61
3' GTCCTCTTGGAGAGCCGAAAGGGCGCGCCGCGGCGGAGAACGACGCGGAGGCGGAGGAGGA 5'
Direction of Transcription
-91

and at positions -91, -90, and -75 (relative to the translation initiation codon) in the human gene (63). Close inspection of the DNA sequences associated with these footprints reveals a perfectly conserved and shared 9 bp sequence of CGCCGCCTC that is apparent when the mouse gene is displayed in the reverse orientation (see Figure 2.5b). Furthermore, the same guanine nucleotide (the second G) is footprinted in both species within this conserved sequence. Thus, this sequence is likely to represent the core binding sequence of a conserved transcription factor that interacts with this region in both genes and activates transcription of the HPRT gene on the active X chromosome. The position of this sequence near the transcription initiation site(s) suggests this factor may act as a transcriptional initiator element (134, 138, 150) for these two TATA-less promoters (see Figure 2.5a). A search of the transcription factor database ((41); and TFSITES at http://ben.vub.ac.be/srs/srsc) yielded no previously identified transcription factor that binds to this conserved 9 bp sequence. Therefore, it is likely that our comparative in vivo footprint analysis of the mouse and human HPRT genes has identified a novel cis-acting regulatory element that interacts with a currently unidentified transcription factor. Curiously, no in vivo footprints are detected at potential E2F/HIP1 binding sites (7, 99, 137) adjacent to the conserved 9 bp sequence in both species (see Figure 2.5).

Genomic Sequencing

LMPCR genomic sequencing was used to examine the methylation status of every cytosine in the 5' region of the mouse HPRT gene on the active and inactive X chromosomes. The region covered was essentially the same region analyzed by in vivo footprinting. Methylation analysis was performed on the 3T3-L1, B5, and C3 cell lines described previously. The same method used to separate the active and inactive mouse HPRT alleles in B5 and C3 cells for in vivo footprinting was also used for methylation analysis. Thus, methylation data from 3T3-L1 DNA and the size-selected DNA of B5 cells yielded the methylation status of the 5' region from the active allele, and size-selected DNA from C3 cells assayed the inactive allele. Because 5-methylcytosine in genomic DNA is more resistant to hydrazine modification than unmethylated cytosine, missing bands in a Maxam and Gilbert cytosine-specific DNA sequencing ladder (relative to an unmethylated control sample) indicates the position of methylated cytosine residues (62). Cytosine residues were considered to be fully methylated when the intensity of a given band was considered to be 25% the intensity of the same band from an unmethylated sample, as estimated by visual inspection. Cytosine residues exhibiting approximately 25% to 80% of the unmethylated band intensity were considered to be partially methylated. If the intensity of a band was greater than 80% the intensity of the unmethylated control band, the specific cytosine site was considered to be unmethylated (61). Because CpG dinucleotides are generally not methylated in *E. coli*, purified plasmid DNA containing the 5' region of the mouse HPRT gene was subjected to genomic sequencing and used as an unmethylated control sample. Eight LMPCR primer sets (see Figure 2.1) were used to analyze the upper and lower strand of the 5' region of the *M. musculus* HPRT gene. Representative autoradiograms from these genomic sequencing studies are shown below and included the GC box region that exhibited the unusual methylation pattern on the inactive human HPRT gene.

FIGURE 2.6 - Genomic sequencing of the upper strand and lower strand using primer sets J and U. A. Autoradiogram of the cytosine-specific sequencing ladder from position +90 to -79 using primer J. B. Autoradiogram of the cytosine-specific sequencing ladder from position -2 to -103 using primer U. The horizontal bars to the left of the sequencing ladders indicate the position of cytosines in CpG dinucleotides. The position relative to the major transcription start site is shown to the right of the sequencing ladder. Bracketed regions with Roman numerals to the right denote the positions of the GC boxes. DNA from the following sources was used for genomic sequencing: lane 1, cloned plasmid DNA containing the 5' region of the mouse HPRT gene; lane 2, purified DNA from the active 5' region of the *M. musculus* HPRT gene (from the B5 cell line); lane 3, purified DNA from the inactive 5' region of the *M. musculus* HPRT gene (from the C3 cell line); and lane 4, purified DNA from 3T3-L1 cells which contain an active mouse X chromosome.

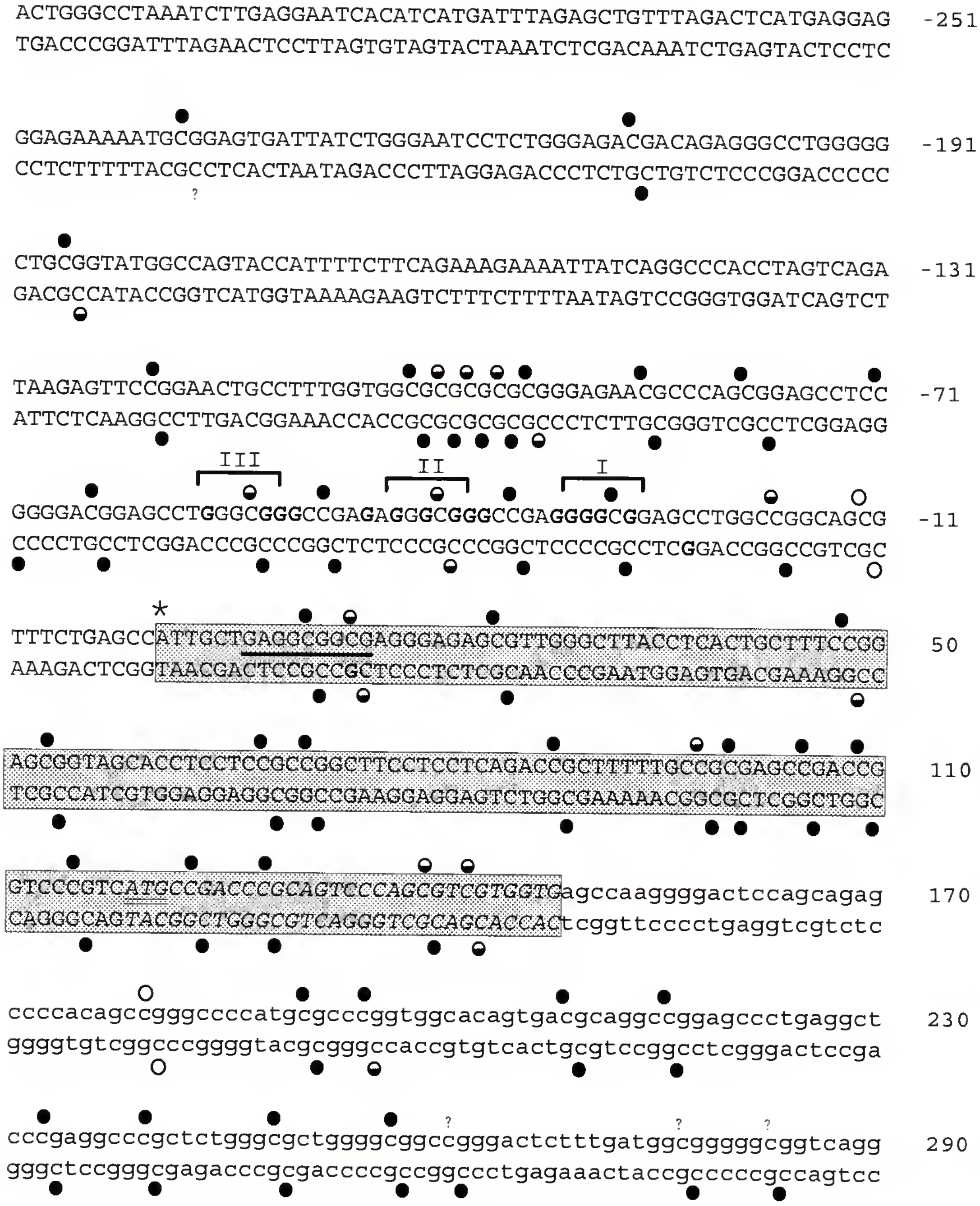

A
1
2
3
4
+90
+25
-17
I
II
III
-79
B
1
2
3
4
-103
-65
III
II
I
-23
-2

In Figure 2.6b, primer set U was used to examine the methylation of cytosines on the upper strand between positions -2 and -103. The region includes the major transcription start site and the three GC boxes. Within this region, all cytosines on the active allele were unmethylated (lanes 2 and 4) since the hydrazine modification and cleavage pattern for the size-selected B5 genomic DNA (lane 2) and 3T3-L1 genomic DNA (lane 4) was identical to the pattern from the unmethylated control plasmid DNA containing the mouse HPRT 5' region (lane 1). On the inactive allele from size-selected C3 genomic DNA (lane 3), the majority of CpG's on the inactive allele were fully methylated. However, a number of CpG's on the inactive allele were unmethylated or partially methylated; for instance, the CpG at -12 just 3' of GC box I was unmethylated and cytosines at positions -18, -41, -54, -98,-99, and -100 on the upper strand were determined to be partially methylated.

The lower strand was examined by primer set J (Figure 2.6a), which covers roughly the same region as primer set U on the opposite strand. Again, all cytosines examined on the active allele were unmethylated (lanes 2 and 4), while most CpG's on the inactive allele were fully methylated with some CpG's either partially methylated or unmethylated (lane 3).

Genomic sequencing was also performed using LMPCR primer sets G, H, I, R, S, and T (data not shown) and data obtained from the inactive allele using all eight primer sets are summarized in Figure 2.7. All cytosines examined in the 5' region of the active allele were unmethylated; thus the methylation pattern of the active mouse and human HPRT genes is highly conserved. On the inactive allele, the majority of CpG dinucleotides in the region were fully methylated and interspersed with a small number of

FIGURE 2.7 - Summary of the methylation pattern from the 5' region of the inactive mouse HPRT gene. Numbering is relative to the major transcription start site as denoted by the asterisk. Bolded nucleotides indicate in vivo footprinted guanine nucleotides (see Fig. 2.4). The shaded box indicates the first exon with italicized letters denoting the translated coding region; the translation start site is indicated with a thin double line. Lower case letters denote the beginning of first intron. The three GC boxes are indicated by brackets with Roman numerals (and correspond to the Roman numerals and brackets in all other Figures). Solid circles indicate methylated cytosines, open circles indicate unmethylated cytosines, and half-filled circles indicated partially methylated cytosines. Question marks denote cytosines for which the methylation status could not be determined. The bold underline beginning at position +7 indicates the position of the conserved footprinted 9 bp sequence that is similarly footprinted in the human HPRT gene (see Fig. 2.5 and text).


ACTGGGCCTAAATCTTGAGGAATCACATCATGATTTAGAGCTGTTTAGACTCATGAGGAG -251
TGACCCGGATTTAGAACTCCTTAGTGTAGTACTAAATCTCGACAAATCTGAGTACTCCTC
GGAGAAAAATGCGGAGTGATTATCTGGGAATCCTCTGGGAGACGACAGAGGGCCTGGGGG -191
CCTCTTTTTACGCCTCACTAATAGACCCTTAGGAGACCCTCTGCTGTCTCCCGGACCCCC
?
CTGCGGTATGGCCAGTACCATTTTCTTCAGAAAGAAAATTATCAGGCCCACCTAGTCAGA -131
GACGCCATACCGGTCATGGTAAAAGAAGTCTTTCTTTTAATAGTCCGGGTGGATCAGTCT
TAAGAGTTCCGGAACTGCCTTTGGTGGCGCGCGCGGGAGAACGCCCAGCGGAGCCTCC -71
ATTCTCAAGGCCTTGACGGAAACCACCGCGCGCGCCCTCTTGCGGGTCGCCTCGGAGG
III
II
I
GGGGACGGAGCCTGGGCGGGCCGAGAGGGCGGGCCGAGGGGCGGAGCCTGGCCGGCAGCG -11
CCCCTGCCTCGGACCCGCCCGGCTCTCCCGCCCGGCTCCCCGCCTCGGACCGGCCGTCGC
*
TTTCTGAGCCATTGCTGAGGCGGCGAGGGAGAGCGTTGGGCTTACCTCACTGCTTTCCGG 50
AAAGACTCGGTAACGACTCCGCCGCTCCCTCTCGCAACCCGAATGGAGTGACGAAAGGCC
AGCGGTAGCACCTCCTCCGCCGGCTTCCTCCTCAGACCGCTTTTTGCCGCGAGCCGACCG 110
TCGCCATCGTGGAGGAGGCGGCCGAAGGAGGAGTCTGGCGAAAAACGGCGCTCGGCTGGC
GTCCCGTCATGCCGACCCGCAGTCCCAGCGTCGTGGTGagccaaggggactccagcagag 170
CAGGGCAGTACGGCTGGGCGTCAGGGTCGCAGCACCACtcggttcccctgaggtcgtctc
ccccacagccgggccccatgcgcccggtggcacagtgacgcaggccggagccctgaggct 230
ggggtgtcggcccggggtacgcgggccaccgtgtcactgcgtccggcctcgggactccga
? ? ?
cccgaggcccgctctgggcgctggggcggccgggactctttgatggcgggggcggtcagg 290
gggctccgggcgagacccgcgaccccgccggccctgagaaactaccgcccccgccagtcc

partially methylated and unmethylated CpG's that appeared to be randomly scattered throughout the region. This is in contrast to the unusual methylation pattern of the inactive human HPRT gene where nearly all of the partially methylated and unmethylated CpG's are concentrated over the cluster of GC boxes (61). Therefore, other than the fact that both the mouse and human genes are hypermethylated on the inactive allele, the most notable aspect of the methylation pattern on the inactive human HPRT gene is not conserved in the mouse gene.

Discussion

In Vivo Footprinting

We have employed a novel modification of the LMPCR method of in vivo footprinting (and genomic sequencing) to separate and analyze the active and inactive mouse HPRT alleles in female somatic cells. This approach, based on the use of species-specific restriction sites in female cells from interspecific crosses, should be useful for similar LMPCR-based studies of other differentially expressed alleles such as imprinted genes, or closely related genes within a gene family. This is, to date, the only non-human X-linked gene analyzed by in vivo footprinting of the active and inactive X chromosomes, and for the first time permits a detailed in vivo comparison of regulatory elements in the promoter region of the murine and human homolgues of an X-linked gene.

Comparison of the in vivo footprint patterns in the 5' region of the mouse and human HPRT genes revealed a strong degree of conservation. Analysis of both genes showed multiple sequence-specific DNA-protein interactions specific only to the transcriptionally active alleles; no footprints were detected in this region of the inactive alleles. The active mouse gene displayed footprints over a cluster of three GC boxes, a sequence known to bind the human transcriptional activator Sp1 (11). Similarly, footprints were observed over a cluster of 5 GC boxes on the active human HPRT gene (63). Another footprinted site near the transcription initiation sites of the active mouse gene is also footprinted in the human gene. The one notable difference between the two genes was a footprinted AP2 site in the human gene that was not detected in the mouse gene.

The footprint on the active mouse gene near the transcription initiation sites (a single site of enhanced DMS reactivity located at position +14; see Fig. 2.4) occurs within a 9 bp sequence that is perfectly conserved and identically footprinted in the human HPRT promoter (see Fig. 2.5). Thus, this 9 bp element is likely to constitute the core binding sequence for a transcriptional activator protein(s). A search of current data bases ((41); and TFSITES at http://ben.vub.ac.be/srs/srsc) revealed no previously described transcription factors bind to this sequence, suggesting that our analysis has identified a new cis-acting regulatory element. The position of this conserved 9 bp sequence, within or adjacent to a region of multiple transcription initiation sites in the mouse and human HPRT genes (101, 115), suggests that it may act as a transcriptional initiator element (150). This sequence does not exhibit homology with the consensus binding sequences of known initiator elements (150).

Further inspection of the region adjacent to the footprint at position +14 of the mouse gene shows a potential binding site for the transcription factor E2F/HIP1 (7, 41, 99, 137) just upstream of the 9 bp conserved sequence, a binding site that is also present just upstream of the same 9 bp sequence in the human HPRT gene. However, the potential E2F/HIP1 site does not exhibit an in vivo DMS footprint in the mouse HPRT gene, nor does it show an in vivo footprint with DMS (63) or DNase I (Y. Yu, T.P. Yang, unpublished data) in the human HPRT gene.

Currently, in vivo footprint analysis has been performed on the 5' region of three genes on the active and inactive X chromosomes: mouse and human HPRT (63), and human PGK-1 (122). Overall, the in vivo footprint patterns are similar; all three genes carry multiple in vivo footprints on the active X chromosome and do not exhibit footprints on the inactive X chromosome. In addition, all three genes contain multiple GC boxes which are footprinted on the active allele. However, aside from GC boxes (which also regulate autosomal genes), no in vivo footprinted DNA sequence common to all three genes has been identified. The conserved 9 base pair sequence associated with the HPRT genes (see above) is absent in the human PGK-1 gene; an AP2 site, footprinted in the human HPRT gene, is absent in both the mouse HPRT gene and the human PGK-1 gene; and a number of footprinted sequences in the PGK-1 gene are not present in the human and mouse HPRT genes (63, 122). These results support the notion that there is no common cis-acting regulatory element in the immediate promoter region that is common to all X-linked genes subject to X inactivation.

Genomic Sequencing

The mechanism by which DNA methylation may silence transcription on the inactive X chromosomes in mammals is unknown. To identify conserved methylation patterns on the inactive allele of X-linked genes, we employed LMPCR genomic sequencing to determine the methylation state of every cytosine nucleotide in the 5' region of the mouse HPRT gene on the active and inactive X chromosomes. These studies produced results generally consistent with previous methylation studies of X-linked genes (61, 71, 89, 122, 143, 144). Every CpG examined on the transcriptionally active mouse HPRT gene is unmethylated, while the methylation pattern of the inactive allele is heterogeneous, with the majority of CpG's fully methylated and interspersed with unmethylated or partially methylated CpG sites (see Fig. 2.7). Similar studies performed previously on the inactive human HPRT gene showed it has an unusual pattern of hypermethylation; all CpG sites in the 5' region are fully methylated except over a region of clustered Sp1 binding sites where most CpG's are either partially methylated or unmethylated (61). Results from the current study of the mouse HPRT gene indicate this unusual pattern of methylation is not conserved on the inactive mouse HPRT gene; the cluster of three GC boxes in the mouse gene is not hypomethylated relative to adjacent regions. Thus, we conclude that the unusual high resolution methylation pattern seen in the inactive human HPRT gene is not conserved on the inactive mouse HPRT gene, and is therefore unlikely to play a crucial role in the differential transcription of the HPRT gene on the active and inactive X chromosomes.

Results of comparable DNA methylation analysis of the mouse HPRT gene by Park and Chapman using the bisulfate genomic sequencing method (114) were generally similar to those obtained in our study. They found all CpG sites in the 5' region of the active allele to be completely unmethylated, and CpG's on the inactive allele to be hypermethylated. However, on the inactive allele, the levels of methylation at individual CpG sites in our studies were not identical to those reported by Park and Chapman from cultured cells. For example, a CpG site at position -12 is unmethylated in our study while Park and Chapman show high levels of methylation at this site in their samples.

Furthermore, we do not observe unusually high levels of methylation within the three hypermethylated "domains" of the promoter reported by Park and Chapman; instead, we find a generally consistent level of CpG methylation on the inactive HPRT allele from positions -114 to +156 (which includes 18 additional CpG sites upstream and downstream of those assayed by Park and Chapman). However, it is not clear that the differences in these two studies (probably due to the use of different cells) are functionally significant since the HPRT gene on the inactive X chromosome in both studies remains transcriptionally repressed.

In fact, examination of the inactive allele of four X-linked genes (human PGK-1, mouse PGK-1, human HPRT, and mouse HPRT; (61, 122, 143)) by genomic sequencing reveals no obviously conserved pattern of methylation in the 5' region that is common to all four genes which might provide an understanding of the mechanism by which methylation functions in repressing transcription of genes on the inactive X chromosome. Furthermore, we do not find any correlation between the methylation patterns of CpG sites and binding sites for transcription factors as determined by in vivo footprinting (63). Aside from the tract of hypomethylation over the GC boxes in the human HPRT gene, high resolution methylation analysis of the 5' region of these four genes reveals no clear and consistent difference between footprinted and non-footprinted regions. Even the region near the transcription initiation site, a region reported to be important for transcriptional silencing by DNA methylation (82, 121), does not show an unusually high density of fully methylated sites on the inactive mouse HPRT and PGK-1 genes (relative to surrounding regions). In fact, one of the few unmethylated sites on the inactive mouse

HPRT gene is found just upstream of the major initiation site at position -11 (see Fig. 2.7).

Implications for X Chromosome Inactivation

These in vivo footprinting studies demonstrate for the first time in the mouse that it is unlikely an X chromosome-specific regulatory sequence(s) within the promoter region of each X-linked gene coordinately regulates transcription of individual genes by X inactivation, consistent with similar investigations of human X-linked genes (63, 122). Also consistent with data from human gene studies (63, 119, 122), there is no evidence for: 1) sequence-specific DNA-protein interactions in the promoter region that repress transcription of genes on the inactive mouse X chromosome; and 2) binding of a methylated DNA binding protein (84) on the inactive allele of the mouse HPRT gene. Thus, sequence- or methylation-specific DNA-protein interactions in the immediate 5' region of individual X-linked genes do not appear to be the fundamental mechanism responsible for maintaining the differential transcription of genes on the active and inactive X chromosomes in somatic cells. Rather, the preferential binding of ubiquitous transcriptional activators to promoter regions on the active X chromosome is more likely to be a secondary effect of differences in chromatin structure between the active and inactive alleles, and/or due to regulatory elements located outside of the immediate promoter region examined by these in vivo footprinting and genomic sequencing studies. The absence of methylation- or sequence-specific protein binding to the inactive 5' region suggests that chromatin structure and/or other more distal regulatory elements may be regulating accessibility of transcription factors to the immediate promoter region on the inactive X chromosome.

The lack of a common conserved methylation pattern in the promoter region of the inactive allele of the human and mouse HPRT and PGK-1 genes (61, 122, 143) - aside from an overall level of hypermethylation relative to the unmethylated active allele - suggests the absence of specific CpG sites or clusters in the immediate 5' region that are common to all of these genes and are crucial for regulation of individual X-linked genes by DNA methylation. In fact, the methylation pattern of the inactive mouse PGK-1 gene, where only a single CpG site (in the 5' untranslated region) is fully methylated, would suggest that crucial sites for regulation of X-linked genes by DNA methylation, if they exist, may be located outside of the immediate promoter of these genes.

CHAPTER 3

ELECTROPHORETIC GEL MOBILITY SHIFT ASSAYS OF A DNA-PROTEIN INTERACTION SPECIFIC TO THE ACTIVE ALLELE OF THE HUMAN AND MOUSE HPRT GENES

Introduction

In vivo footprinting analysis of the human and mouse HPRT 5' regions revealed several footprinted sites on the active X chromosomes. Of the footprints observed on the human and mouse HPRT gene, only one appeared to have occurred over a novel sequence not previously reported to bind a transcription factor (63, 87). The remaining footprints occurred over potential Sp1 binding sites (11) or an AP-2 binding site (63) (no AP-2 binding sites were observed in the mouse HPRT gene). The novel in vivo footprint was characterized by a strong site of DMS modification at position -91 and +14 in the human and mouse HPRT genes, respectively. These positions are located just 3' of the transcription start site in both genes. Comparison of the mouse and human HPRT sequences within this region revealed a 9-bp sequence which is conserved in both the human and mouse genes (Figure 2.5). In addition, the same guanine residue in these sequences shows enhanced DMS modification in both the human and mouse HPRT genes.

A search of several transcription factor data bases on the World Wide Web using the 9-bp sequence did not reveal identity or significant homology to any previously described cis-acting regulatory elements among eukaryotic DNA control sequences (41). The fact that this sequence is located just 3' of the multiple transcription start sites in the human and mouse HPRT genes suggests the protein(s) binding at this site may function as a transcriptional initiator. However, there appears to be no sequence similarity between known transcriptional initiators such as Inr, YY1, and TFII-I (134, 138, 150) in the 9-bp sequence. This suggests that the DNA-protein interaction at position -91 and +14 in the mouse and human HPRT may represent a new regulatory element involved in transcription initiation.

Previously, a 104 bp Bsu36I/BssHII fragment which contains only the -91 footprint and the unique 9-bp sequence was characterized for DNA-protein interaction by electrophoretic gel mobility shift assays (59). Reconstitution of the DNA-protein interaction in vitro using crude Hela nuclear extracts revealed multiple sequence-specific DNA-protein complexes. At least four complexes (I-IV) appeared to be specific. These complexes were able to be specifically competed away by the human HPRT, mouse HPRT, mouse DHFR, and mouse APRT promoters. These promoters are all GC-rich housekeeping promoters that lack TATA boxes. The human factor VIIIC and albumin tissue-specific promoters did not significantly compete complexes I-IV. These data suggest that complexes I-IV are specific for the human HPRT 104 bp fragment.

Due to the complexity of the DNA-protein interactions using the 104 bp fragment, it is difficult to determine if these interactions are specific for the -91 footprint, represent other distinct DNA-protein interactions, or represent protein-protein interactions. Therefore, I employed electrophoretic gel mobility shift (EMSA) assays to reconstitute the DNA-protein interactions using crude Hela nuclear extract and small double stranded oligos which span the -91 footprint. These studies should further simplify and characterize the DNA-protein interactions within this region. Furthermore, these experiments may help define the role of this unique binding site in the regulation of HPRT gene expression as well as provide insight in the transcription initiation of TATA-less genes. These experiments are crucial for further characterization and identification of the protein(s) that interacts at the 9-bp sequence.

Material and Methods

Nuclear Extracts

Nuclear extracts were purchased form Gibco BRL. These extracts were prepared as described by Dignam et al. (37).

Preparation of Double Stranded Oligos for Gel Mobility-Shift Assasys

In order to generate 5 doubled-stranded oligos with 5' overhangs that spanned the -91 footprint (See Figure 3.1), ten oligos were synthesized using the human HPRT sequence. The name and nucleotide sequence of the oligos are as follows: 195, CGACGAGCCCTCAGGCGAACCTCTCGGCTT; 196, GGAAAGCCGAGAGGTTCGCCTGAGGGCTCG; 197 GAGGCGCAGCAAGAGGCGGCGCCGCGCG; 198 TCCCGCGCGGCGCCGCCTCTTGCTGCGC; 201 GGCGGCGCCGCGCGGGAAAGCCGAGAGG; 202 GAACCTCTCGGCTTTCCCGCGCGGCGCC; 214, GCCTCTTGCTGCGCCTCCGCCTCCTCCT; 215, CAGAGGAGGAGGCGGAGGCGCAGCAAGA; 216 AGCCGGTGGCGGAGCAGAGGAGGAGGCG; and 217, CTCCGCCTCCTCCTCTGCTCCGCCACCG. 50 pmol of each set of complementary oligos were annealed in a thermal cycler (Coy II) at 95°C for 2 minutes then ramped slowly down for 180 minutes to 2°C. Each double-stranded oligo was radiolabelled with ^{32}P-α-dCTP using the Klenow fragment of DNA polymerase I to fill the 5' overhang (20). Labeled oligos were run on a non-denaturing 10% (80:1 acrylamide:bis-acrylamide) electrophoresis gel in 0.5 X TBE at 400 volts. The position of labeled oligos were located on the gel by exposure to type 57 polaroid film for 10 minutes, excised from the gel with a razor, and purified using the Qiaex II gel purification kit (Qiagen). Purified labeled ds-oligos were diluted to 100,000 counts per minute per micoliter.

Preparation of Sequence-Specific Competitors

An 111 base pair fragment (-237 to –124), 120 base pair fragment (-19 to 102), and 104 base pair fragment (-134 to –16) were prepared by PCR amplification (See Figure 3.1). Primers GTAGGCGCCGGCGTAGGCGC and GGCTCGTCGCAGCCGCCTGA were used to amplify a 111 base pair fragment which contains the 5 potential Sp1 binding sites but does not contain the novel –91 binding site or the AP-2; primers CGCGCCGGCCGGCTCCGTTA and CCCGCGCCTCAGGGCCCGGC were used to amplify a 120 base pair fragment which is 3' of the novel binding site and contains no in vivo footprints, and primers TGCGACGAGCCCTCAGGCGA and CGCGCGGGCTGACTGCTCAG were used to amplify a 104 base pair fragment which contains the -91 footprint and was previously characterized by EMSA. For PCR, 50 ng of plasmid DNA containing 1.8 kb of the human HPRT 5' region (101, 115) were used as a template in 100 ul reaction containing 1.5 mM $MgCl_2$, 0.1 mM each primer, 100 mM dNTP's, and 1 unit Taq polymerase. Reactions were denatured at 95°C for 20 seconds, annealed at 64°C for 30 seconds, and extended at 76°C for 1 minute for 30 cycles in an MJ-100 thermocycler. Following

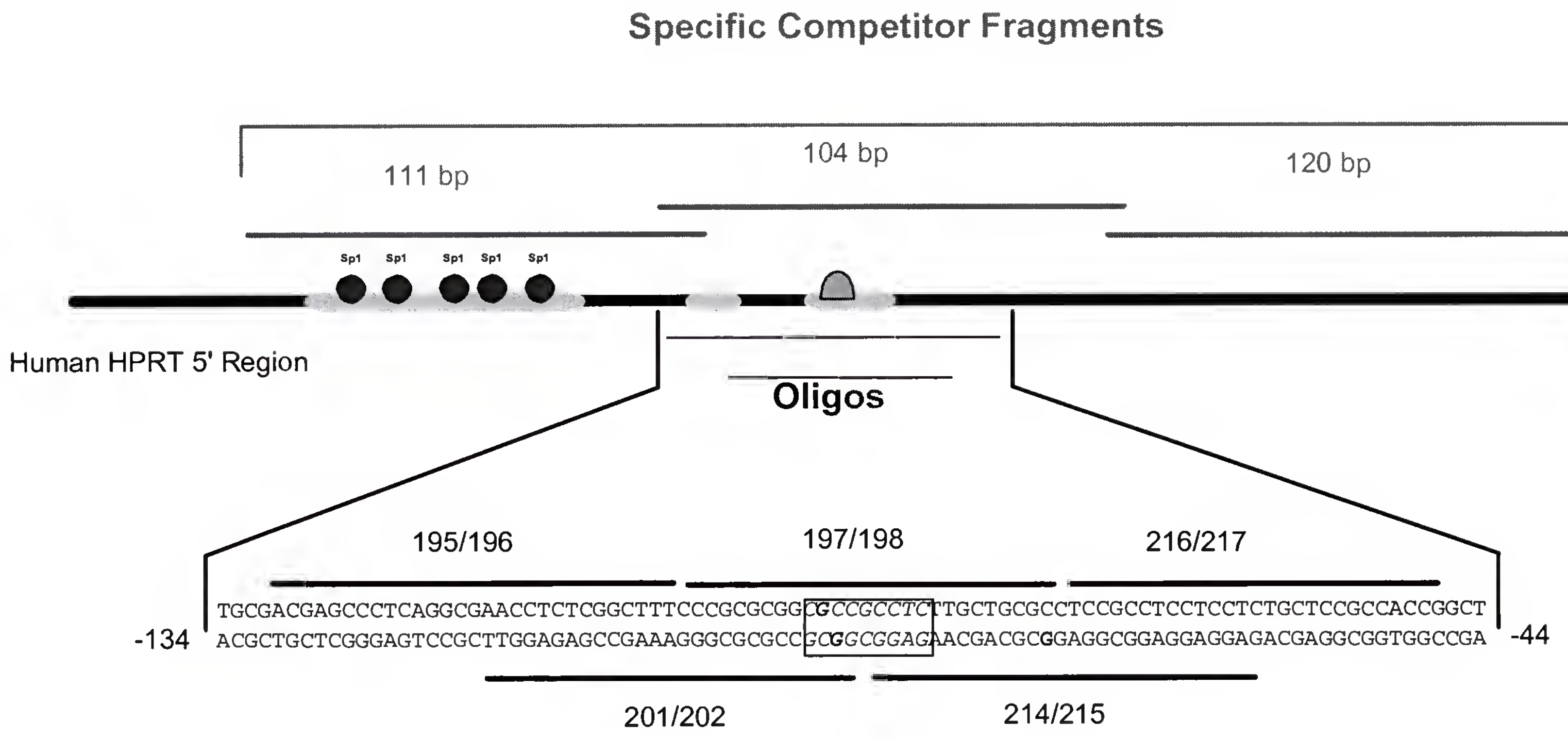

Specific Competitor Fragments
111 bp
104 bp
120 bp
Sp1
Sp1
Sp1
Sp1
Sp1
Human HPRT 5' Region
Oligos
195/196
197/198
216/217
-134
TGCGACGAGCCCTCAGGCGAACCTCTCGGCTTTCCCGCGCGGCGCCGCCTCTTGCTGCGCCTCCGCCTCCTCCTCTGCTCCGCCACCGGCT
ACGCTGCTCGGGAGTCCGCTTGGAGAGCCGAAAGGGCGCGCCGCGGCGGAGAACGACGCGGAGGCGGAGGAGGAGACGAGGCGGTGGCCGA
-44
201/202
214/215

FIGURE 3.1 - Diagram of the oligos and sequence-specific competitors used in gel mobility-shift assays. The human HPRT sequence is designated by the thick black line. The shaded ovals designate DNaseI in vivo footprints detected on the active human HPRT gene. DMS in vivo footprints are indicated by shaded circles and half circles. The expanding bracket indicates the sequence and region of the human HPRT gene examined with oligos in the gel mobility shift assays. The box indicates the novel nine base sequence that is identical in both the human and mouse HPRT genes. Bold letters designate DMS in vivo footprints detected on the active human HPRT gene. The thick lines labeled 195/196, 197/198, 216/217, 201/202, 214/215 indicate the positions of the oligos used in the gel mobility-shift assays. The thick gray lines labeled 111 bp fragment, 104 bp fragment, and 120 bp fragment indicate the positions of the sequence-specific competitors fragments.

amplification, fragments were size fractionated on a 1.8 % agarose gel in 0.5 X TBE, excised from the gel, and purified using the Qiaex II gel purification kit (Qiagen).

The 1.8 kb EcoRI-BamHI human HPRT fragment was prepared from the pl4X8-RB1.8 plasmid. Approximately 500 mg of the pl4X8-RB1.8 plasmid was digested with 1000 units of EcoRI and BamHI. The fragment was size fractionated on a 0.8% agarose gel in 0.5 X TBE, excised from the gel and purified using the Qiaex II gel purification kit. All purified fragments were then quantitated and diluted to specific molar concentrations.

Electrophoretic Gel Mobility-Shift Assays

The 10 ul binding reactions (20) consisted of 100,000 counts per minute of labeled fragment (labeled oligos 201/202, 197/198, and 195/196), 1 ug dI-dC, dA-dT, or sonicated calf thymus DNA as nonspecific competitor, 2 ug of crude Hela nuclear extract, in 1X binding buffer (12% glycerol, 12 mM HEPES NaOH, pH 7.9, 60 mM KCL, 5 mM $MgCl_2$, 4 mM Tris, pH 8, 0.6 mM EDTA, 0.6 mM dithiothreitol). First, buffer, water, nuclear extract, and nonspecific competitor were mixed and allowed to incubate 5 minutes at 25°C. Next, specific competitor was added and the reaction was incubated 5 minutes at 25°C. Finally, labeled fragment was added and the reaction was allowed to incubate 20 minutes at 25°C. Following the last incubation, the binding reaction was size-fractionated on a 6% acrylamide (80:1 acrylamide:bis-acrylamide) gel containing 0.5 X TBE. Better size-fractionation of DNA-protein complexes was obtained with 0.5 X TBE than low ionic strength gel buffer. After electrophoresis, the gel was dried and exposed to Kodak XAR film for 1-3 days or Kodak Biomax film for 1 day.

Results

Comparison of the mouse and human HPRT sequence revealed a 9-bp sequence which was identical in both genes and in vivo footprinted at the same guanine residue (-91 in the human HPRT gene and +14 in the mouse HPRT gene) (63, 87). To reconstitute in vitro the DNA-protein interaction(s) occurring over this sequence, several double stranded oligos that spanned and flanked this sequence were prepared from the human HPRT sequence. The 197/198 double-stranded (ds) oligo spans the –91 in vivo footprint observed on the active human HPRT gene (63). The 195/196 ds-oligo is 5' of the -91 footprint and spans a potential DNase I footprint [C. Chen, Y. Yu and T.P. Yang, in preparation]. The 201/202 ds-oligo is located between the -91 footprint and the DNase I footprint. The 216/217 ds-oligo is 3' of the -91 in vivo footprint and spans no observed footprints. These four ds-oligos do not contain any sequence from the GC box region that was footprinted in vivo on the active human HPRT gene (63). The locations of the oligos relative to the DMS and DNase I in vivo footprints are shown in Figure 3.1. Electrophoretic gel mobility-shift assays were used to detect the binding of sequence-specific transcription factors to each double-stranded oligo (43, 44). Labeled ds-oligos were prepared as described in the materials and methods, incubated with crude Hela nuclear extracts (37), and the resulting DNA-protein complexes were size fractionated on a native acrylamide gel.

Figure 3.2 shows the results of the gel mobility-shift assays for the 195/196, 197/198, 201/202, and 216/217 oligos of the human HPRT promoter. When these oligos are incubated with Hela nuclear extract, proteins bind to the DNA fragments. These

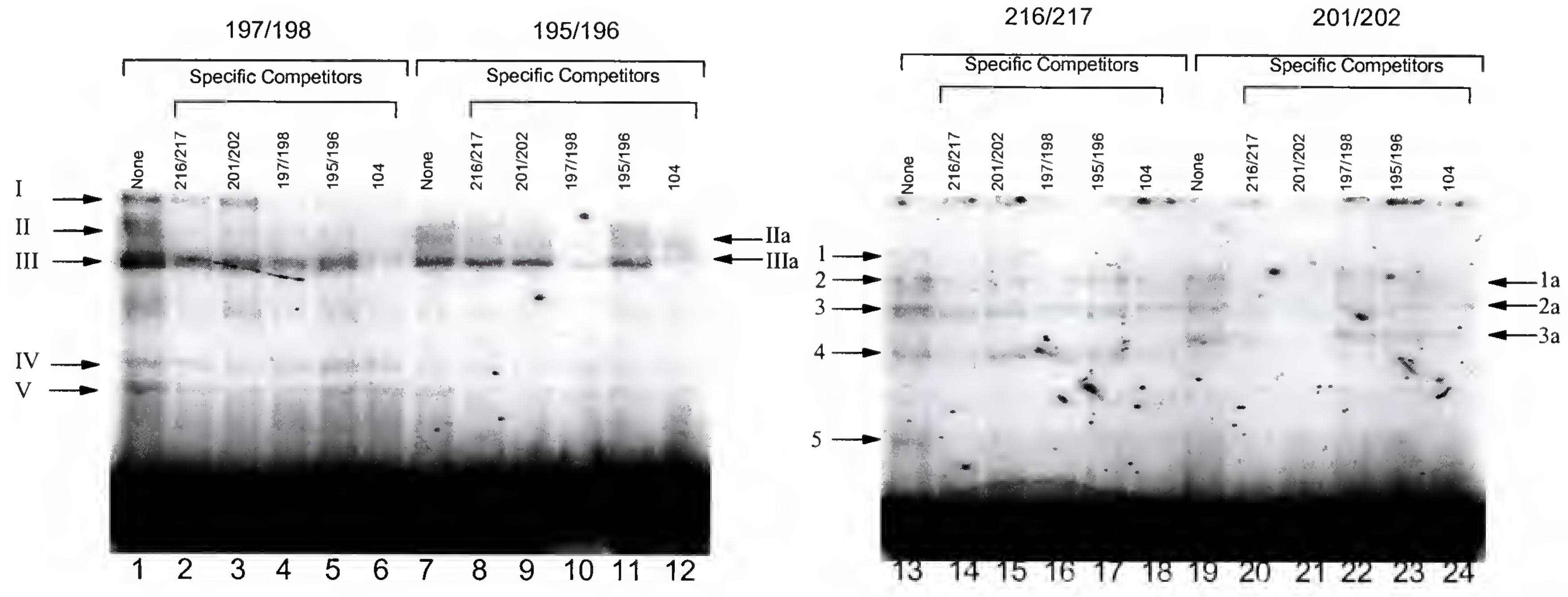


FIGURE 3.2 - Electrophoretic mobility-shift assays using the 195/196, 197/198, 201/202, and 216/217 double-stranded oligos as probes. The lanes are numbered at the bottom. All samples were incubated with 2.5 ug of Hela nuclear extract and 0.75 ug of the nonspecific competitor dI-dC. The brackets labeled 197/198, 195/196, 216/217, and 201/202 indicate samples in which that oligo was used as a probe. None denotes reactions with no specific competitor. 216/217, 201/202, 197/198, and 195/196 indicate the double-stranded sequence-specific oligo used in 1000 molar excess as a competitor. 104 indicates the 104 bp sequence-specific fragment was used as the competitor in 100 molar excess. The roman numerals and arrows indicate potential specific DNA-protein complexes observed.

DNA-protein complexes can be visualized as bands with retarded mobility in the autoradiogram. The amount and type of nonspecific competitors and Hela nuclear extract were optimized in previous experiments (data not shown). In these experiments, when nuclear extract concentration was increased, an increase in the molecular weight and number of DNA-protein complexes was observed. The nonspecific competitors dA-dT, dI-dC, and calf thymus DNA were titrated from 0.25 ug to 2 ug to minimize the formation of nonspecific complexes. These experiments indicate that 2.5 ug of Hela nuclear extract and 0.75 ug of the nonspecific competitor dI-dC were optimal.

In Figure 3.2, lanes 1, 7, 13, and 19 show the DNA-protein complexes formed in the absence of sequence-specific competitors for the ds-oligos 197/198, 195/196, 201/202, and 216/217, respectively. Sequence-specific DNA competitors from the human HPRT sequence were used to determine which retarded bands represent sequence-specific DNA-protein interactions. Lanes 2-6, 8-12, 14-18, and 20-24 show the DNA-protein complexes formed with the four different ds-oligos (197/198, 195/196, 216/217, and 201/202, respectively) after the addition of Hela nuclear extract in the presence of five different sequence-specific competitors. These competitors, which were added at 1000 fold molar excess except the 104 bp fragment which was added at 100 fold molar excess, are the 216/217 ds-oligo (lanes 2, 8, 14, and 20), the 201/202 ds-oligo (lanes 3, 9, 15, and 21), the 197/198 ds-oligo (lanes 4, 10, 16, and 22), the 195/196 ds-oligo (lanes 5, 11, 17, 23), and the 104 bp fragment (lanes 6, 12, 18 and 24). The positions of the specific competitors are shown in Figure 3.1.

Specific DNA-protein complexes may be indicated by retarded bands that disappear when competed with a sequence-specific competitor. Multiple DNA-protein complexes (denoted as complexes I, II, III, IV, V, Ia, IIa, IIIa, 1, 2, 3, 4, 5, 1a, 2a, and 3a) were observed for the 197/198, 195/196, 216/217, and 201/202 ds-oligos. The 216/217 and the 201/202 ds-oligos were examined by EMSA as controls to help indicate specific DNA-protein interactions. The 216/217 ds-oligo is located 3' of the -91 footprint and spans a region in which no DNase I or DMS in vivo footprints were observed on the active HPRT allele. For the 216/217 ds-oligo (lanes 13-18) 5 potential DNA-protein complexes were observed in the absence of competitor. Of these complexes, complexes 2 and 4 appeared to be competed away by the 216/217 ds-oligo itself (lane 14) and the 104 bp fragment which contains the 216/217 ds-oligo sequence. Complex 1 appeared to be competed away by itself (lane 14), the 197/198 ds-oligo (lane 16), which contains the novel 9-bp sequence, the 195/196 ds-oligo (lane 17), which contains a DNase I footprint 5' of -91, and the 104 bp fragment (lane 18). Complex 5 was competed away by all the sequence-specific competitors (lanes 14-18). Complex 3 was not competed away by any of the sequence-specific competitors, suggesting this complex is nonspecific. These data suggest that complexes 2 and 4 may be specific for sequences in the 216/217 ds-oligo, that complex 1 may be specific for sequence in the 197/198 and 195/196 ds-oligos, and complex 5 is present in all of the competitors. The 201/202 ds-oligo is located between the DNase I footprint and the novel 9-bp sequence. For this ds-oligo 3 DNA-protein complexes were observed in the absence of sequence-specific competitors (lane 19). Complex 3a appeared to be competed away only by the 201/202 ds-oligo which is itself (lane 21) and the 104 bp fragment (lane 24) which contains the sequence of the 201/202 ds-oligo. Complex 1a appeared to be competed away by itself (lane 21), the 216/217 ds-oligo (lane 20) and the 104 bp fragment (lane 24). Complex 2a appeared to be nonspecific, as it was not competed away by any sequence-specific competitors. The 197/198 and 195/195 ds-oligos which spanned different observed in vivo footprints did not appear to compete away any of the complexes shifted by the 201/202 ds-oligos. These data suggest that complex 3a is specific for sequences in the 201/202 ds-oligo, that complex 1 may interact with sequence present in both the 201/202 and 216/217 ds-oligos, and that no complexes are specific for sequences common to the 201/202 ds-oligos, 197/198 ds-oligos and the 195/196 ds-oligos. It should be noted that the potential DNA-protein complexes observed for the 216/217 and 201/202 ds-oligos were faint and disappeared at higher concentrations of the nonspecific competitor dI-dC.

Because the 195/196 and 197/198 ds-oligos spanned different region that were footprinted in vivo, these ds-oligos were examined using EMSA to attempt reconstitution DNA-protein interaction in vitro, which may be indicative of these in vivo footprints. The 195/196 ds-oligo is 5’ of the novel 9-bp sequence and spans a DNase I in vivo footprint observed on the active HPRT allele (Figure 3.1). For the 195/196 ds-oligo, two complexes (complexes IIa and IIIa) were observed in the absence of sequence-specific competitors. Both of these complexes were competed away by the 104 bp fragment which contains the 195/196 sequence as well as both in vivo footprints. However, neither these complexes were competed away by itself (the 195/196 ds-oligo, lane 11) or the other sequence-specific ds-oligos (lanes 8, 9, and 10). This suggests that none of these complexes are specific for sequences in the 195/196 ds-oligo. These experiments should be repeated in order to validate this result. The 197/198 ds-oligo spans the -91 DMS in vivo footprint that occurs within the novel 9-bp sequence. For this ds-oligo five potential DNA-protein complexes were observed in the absence of sequence-specific competitors (lane 1). Complexes IV and V were not competed away by any sequence-specific competitors (lanes 2-6), suggesting this complex is nonspecific. Complex II was competed away by all the sequence-specific competitor (lanes 2-6), suggesting this complex interacts with a sequence which is contained in all the sequence-specific competitors. Complex I appeared to be competed away with the 197/198 ds-oligos which is itself and spans the -91 in vivo footprints (lane 4), the 195/196 ds-oligos which spans the DNase I footprint observed in vivo 5' of the -91 footprints (lane 5), and the 104 bp fragment which spans both in vivo footprints (lane 6). This suggests that complex I interacts with specific sequences common to both the 197/198 and the 195/196 ds-oligo. Because both of these ds-oligo span observed in vivo footprints, it is possible that complex I represents the in vitro reconstitution of the -91 DMS and DNase I in vivo footprints observed on the active human HPRT gene. For the 195/196 ds-oligos a DNA-protein complex of the same size was not observed (lanes 7-12). Complex III appeared to be competed away by only the 104 bp fragment (lane 7) and to a lesser extent by itself (the 197/198 ds-oligo, lane 6). It is possible that complex III represents a sequence-specific complex that interacts with sequences in the 197/198 ds-oligo. A similar DNA-protein complex that shifts at the same position (complex IIIa) is observed for the 195/196 ds-oligo. However, for the 195/196 ds-oligo the complex appears more like a single band and does not appear to compete with itself (lane 11).

Discussion

These in vitro reconstitution studies have demonstrated the formation of multiple DNA-protein complexes on the 197/198, 195/196, 216/217, and 201/202 ds-oligos in the presence of crude Hela nuclear extract. The fact that these complexes are not efficiently competed away by different nonspecific competitors suggests that proteins are specifically binding to sequences within the oligos. Although several DNA-protein complexes were observed for the 201/202 and 216/217 ds-oligos in the presence of extract, the fact that these complexes were barely detectable and were competed away in the presence of 1 ug of dI-dC suggests that they are either nonspecific, or may represent DNA-protein interactions with only a small portion of these ds-oligos sequence. Since the 201/202 ds-oligo overlaps approximately two nucleotides with the DNase I footprint just 5' of the -91 footprint, and the 216/217 ds-oligos overlaps approximately 2 nucleotides with the 3' region of the -91 footprint, it is possible that some of these complexes (such as complex 1 for the 216/217 ds-oligo) may represent in vitro reconstitution of the –91footprint.

The 197/198 oligo spans a novel 9-bp sequence which is conserved in both the human and mouse HPRT gene. In addition, this 9-bp sequence has been shown to be footprinted at the same guanine residue (-91 and +14 in the human and mouse HPRT genes respectively) (63, 87). Furthermore Yongjia Yu et al. (personal communication) observed a DNase I in vivo footprint over this sequence on the active human HPRT gene. Therefore any specific DNA-protein interactions observed on the 197/198 are likely to represent the factor(s) which are interacting with the -91guanine residue in vivo. Of the five DNA-protein complexes observed, complex I and III were able to be competed away by specific competitors (the 104 bp fragment, the 195/196 ds-oligo, and the itself) that either contained the novel nine base pair sequence or the DNase I in vivo footprint just 5' of the -91 footprint (Figure 3.2 lane 6). Other sequence-specific competitors which did not span observed in vivo footprints failed to compete away any of complexes I or III

(Figure 3.2, lanes 2 and 3). This suggests that complex I and III may represent the DNA-protein interactions observed in the regions form -75 to -97 and -113 to -127 by DNase I in vivo footprinting experiments of the active human HPRT 5' region. The fact that complex I appears to be efficiently competed away by itself and the 195/196 ds-oligo suggests that this DNA-protein complex may be interacting with the same sequence in both the 197/198 ds-oligo and the 195/196 ds-oligo. However, a DNA-protein complex of similar size was not observed for 195/196 ds-oligo. It is possible that this DNA-protein complex interacts more efficiently with the region immediately surrounding the -91 rather than the DNase I footprinted region from -113 to -127. This would not be inconsistent with the in vivo footprinting data which showed stronger footprints at the -91 region and would explain the lack of detectable complex of the same size observed in the 195/196 ds-oligo.

The 195/196 oligo is located immediately 5' of the 197/198 ds-oligo. No DMS in vivo footprints were observed in the region covered by the 195/196 ds-oligo. However, DNase I in vivo footprinting experiments suggest that this region (-113 to -127) is bound by a factor on the active human HPRT gene (C. Chen, Y. Yu, T.P.Yang, in preparation). Consequently, the specific DNA-protein interactions observed on the 195/196 oligo may represent a factor which is bound to this region in vivo. For the 195/196 ds-oligo, two DNA-protein complexes were observed in the gel mobility shift assays. The addition of the 104 bp fragment which contains both the -91 footprint and the DNase I footprint observed 5' of the -91 footprint competed away both complexes IIa and IIIa. Complex IIIa a DNA-protein complex of the sample molecular weight was observed for the 197/198 ds-oligos (lanes 1-6, complex III). Furthermore, complex IIIa and III are both competed away by the 104 bp fragment. This suggests that complex IIIa may represent the DNA-protein interaction observed in the DNase I in vivo footprinting experiments. Therefore, complex IIIa may represent a transcription factor that binds just 5' of -91 and at -91 guanine residues on the active human HPRT gene. The simplest explanation for these results is that the same factor is binding to both sites. This is not unlikely because the human HPRT gene contains multiple transcription initiation start sites (116). A second explanation is that two different factors bind these sites in vivo. However, the fact that complex III and complex IIIa are not efficiently competed away by the 197/198 ds-oligo or the 195/196 ds-oligo argues against the latter hypothesis.

Presently, these experiments are preliminary and need to be repeated to demonstrate reproducibility.

CHAPTER 4

5-AZADEOXYCYTIDINE-INDUCED CHROMATIN REMODELING OF THE INACTIVE X-LINKED HPRT GENE PROMOTER OCCURS PRIOR TO TRANSCRIPTION FACTOR BINDING AND GENE REACTIVATION

Introduction

A unique system of differential gene expression in mammals is established during female embryogenesis by X chromosome inactivation (46, 49). The inactivation of one X chromosome within each female somatic nucleus generates a transcriptionally active and inactive allele of most X-linked genes and results in dosage compensation for X-linked genes between males and females. A variety of molecular mechanisms have been implicated in regulating the initiation, spreading, and maintenance of X inactivation (46, 49, 66, 72, 94, 107, 119). The involvement of DNA methylation in this process has been established by studies using methyl-sensitive restriction enzymes (71, 89, 151), DNA-mediated transformation (86, 147, 148), genomic sequencing (63, 122, 143), and the DNA demethylating agent, 5-azacytidine (52, 53, 107, 144, 148). All of these studies support the notion that hypermethylation of the 5' CpG island associated with many X-linked housekeeping genes is involved in the transcriptional silencing of these genes on the inactive X chromosome.

The ability to demethylate and reactivate individual genes on the human inactive X chromosome in rodent-human somatic cell hybrids by treatment with 5-azacytidine (5aC) or 5-aza-2'-deoxycytidine (5aCdr) (40, 107) suggests that transcriptional regulation of X-linked genes by X chromosome inactivation involves some measure of local control either at the level of individual genes or at the level of chromatin domains. Reactivation of inactive X-linked genes such as the hypoxanthine phosphoribosyltransferase (HPRT) and phosphoglycerate kinase (PGK-1) genes following 5aC or 5aCdr treatment is associated with both a change in chromatin structure from a nuclease-inaccessible to a nuclease-accessible conformation, as well as a reduction in DNA methylation levels in the 5' CpG island (40, 152).

In previous studies, Sasaki et al. (133) assayed four parameters during 5aCdr reactivation of the human HPRT gene in a hamster-human somatic cell hybrid cell line (X8-6T2) containing the inactive human X chromosome. The parameters examined were HPRT mRNA levels and three properties of the 5' region, including hemi- and symmetrical demethylation of DNA, and MspI nuclease sensitivity of chromatin. Hemi-demethylation and MspI sensitivity were detectable 6 hr after the addition of 5aCdr and reached maximum levels at 24 hr, while symmetrical demethylation and HPRT mRNA levels became detectable at 24 hr and reached maximum levels 48 hr after exposure to 5aCdr. Thus, the initial events during reactivation of the HPRT gene by 5aCdr treatment are the hemi-demethylation and alteration of chromatin structure in the promoter region, followed by symmetrical demethylation and transcription of the gene. A similar sequence of events is reported for 5aCdr-mediated reactivation of the mouse APRT gene (35).

The major question we address here is whether binding of transcription factors to the promoter region upon 5aCdr reactivation is correlated with the early change in chromatin structure or with actual transcription of the gene (i.e., appearance of mRNA). A recent study of the human hsp70 gene revealed that transcriptional repression at the time of mitosis was associated with a loss of all sequence-specific transcription factors from the promoter region, yet nuclease hypersensitivity was retained (92). Furthermore, Lee and Garrard (81) found that abolishing transcription factor binding to the yeast HSP82 gene promoter did not eliminate a DNase I hypersensitive site within the promoter. This lack of dependence of a nuclease hypersensitive chromatin structure on transcription factor binding in vivo suggests that transcription factor binding to the HPRT promoter following 5aCdr treatment may not occur at the time of nuclease sensitivity acquisition, but rather at a later time that is associated with the onset of transcription.

Analysis by in vivo footprinting shows that the promoters of the active HPRT (63) and PGK-1 (119, 122) alleles are bound by transcription factors, whereas the promoters of the corresponding inactive alleles are devoid of these factors. On the active human HPRT allele, in vivo footprints are associated with each of five potential Sp1 binding sites, a potential AP2 binding site, and a region near the multiple transcription initiation sites (63). For both the HPRT and PGK-1 genes, no evidence has been found for the binding of sequence-specific repressors to the promoters of the inactive alleles. Furthermore, there is no evidence for the interaction of methylated DNA binding proteins (84) with the 5' regions of these genes on the inactive X chromosome. These in vivo footprinting studies indicate that a major component of transcriptional silencing on the inactive X chromosome is the exclusion of transcription factors from promoter regions.

To determine the timing of transcription factor binding during 5aCdr-induced reactivation of the human HPRT gene on the inactive X chromosome, we have performed dimethyl sulfate (DMS) in vivo footprinting on X8-6T2 cells at various times after initiating 5aCdr treatment. We now demonstrate that the binding of transcription factors to three sites in the promoter region correlates with the appearance of HPRT mRNA rather than with the preceding change in nuclease sensitivity of chromatin. Thus, remodeling of chromatin structure during 5aCdr reactivation of the HPRT gene on the inactive X chromosome precedes, and thus, does not require, the binding of at least three sequence-specific transcription factors to the promoter region.

Materials and Methods

Cell Lines

DNA samples were prepared from cultures of cell lines previously described (60, 63). Briefly, GM00468 (NIGMS Human Genetic Mutant Cell Repository) is a normal diploid human male fibroblast cell line containing an active X chromosome. Cell line 4.12 (generously provided by David Ledbetter) is a hamster-human somatic cell hybrid containing only the active human X chromosome in the HPRT-deficient hamster cell line RJK88; RJK88 carries a deletion of the endogenous hamster HPRT gene. GM06318 is also a human-rodent somatic cell hybrid containing an active human X chromosome. Cell lines X8-6T2 and 8121 are hamster-human somatic hybrids containing an inactive human X chromosome (38, 40, 110).

5-Aza-2'-deoxycytidine (5aCdr) Treatment

X8-6T2 cells were treated with 5aCdr as previously described by Sasaki et al. (133). Briefly, cells were treated with 0.4 ug/ml 5aCdr in growth medium (RPMI 1640 with 10% fetal bovine serum and 40 ug/ml gentamicin) for 24 hr. Cells were then washed with phosphate buffered saline (PBS) and returned to normal medium.

MspI Treatment of Chromatin

Assaying chromatin structure changes during 5aCdr reactivation was performed by nuclease digestion of isolated nuclei as described by Sasaki et al. using the restriction enzyme MspI (133). Briefly, nuclei were isolated from X8-6T2 cells 0, 12, 24, 32, 48, and 60 hrs after initiating treatment with 5aCdr. Nuclei from each time point were treated with 0, 200, and 600 units/ml of MspI, genomic DNA was isolated and digested to completion with PstI, then subjected to Southern blot analysis with human HPRT hybridization probe PB1.7. The approximate positions of the clustered MspI sites, the PB1.7 hybridization probe, the 1.6 kb Msp-3'Pst subband, and the transcription factor binding sites are shown in Figure 4.1.

Quantitation of hybridization signals in Southern blots was performed by PhosphorImager analysis (Molecular Dynamics) in the PhophorImager Analysis Facility of the Markey Molecular Medicine Center at the University of Washington. The relative levels of nuclease digestion were calculated as follows. The radioactivity in each 3.3 kb 5'Pst-3'Pst band and 1.6 kb Msp-3'Pst sub-band was quantitatied by PhosphorImager analysis and expressed in PhosphorImager units. Background for the 1.6 kb Msp-3'Pst subfragment was determined by the value of PhosphorImager units at the position of 1.6 kb in samples treated in vivo with 0 units of Msp I. This background value was subtracted from the PhosphorImager units of the 1.6 kb Msp-3'Pst subfragment of each sample treated with MspI to obtain net PhosphorImager units for each Msp-3'Pst sub-band. To quantitate the relative amount of the Msp-3'Pst sub-band at a given time point, the net PhosphorImager units for the band were divided by the total hybridization signal for that sample as defined by the sum of the net units for the Msp-3'Pst sub-band with the

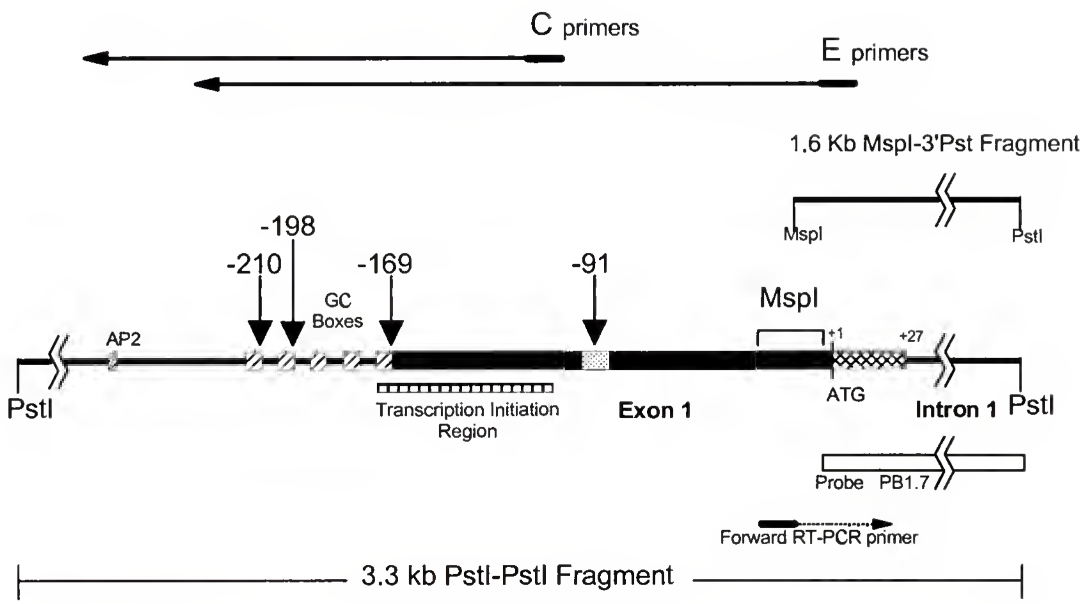


FIGURE 4.1 - Summary diagram of footprints, PCR primers, probes, and DNA fragments used to analyze the human HPRT gene. The long horizontal line indicates the 3.3 kb PstI-PstI restriction fragment which contains the human HPRT promoter, first exon, and first intron. The line is numbered relative to the translation start site which is designated by +1 and ATG. The thick horizontal portion represents the first exon, with the solid portion indicating the 5' untranslated region and the cross-hatched portion indicating the translated portion. The region of multiple transcription initiation sites in the human HPRT gene is designated by a checkered bar below the first exon. The small striped boxes show the positions of the 5 GC boxes, the small shaded box indicates the putative AP2 site, and the small stippled box indicates a novel factor binding site just downstream of the transcription initiation region. The horizontal arrows denoted C primers and E primers indicate the location, direction, and approximate region analyzed by the two LMPCR primer sets. The numbered vertical arrows (-91, -169, -198, -210) indicate the positions of the in vivo footprinted sites examined. The horizontal arrow designated Forward RT-PCR primer indicates the location of the forward primer used for RT-PCR. MspI indicates the position of the 4 clustered MspI sites which where examined in the nuclease-sensitivity studies. The small broken horizontal line designated 1.6 kb MspI-PstI fragment indicates the primary MspI digestion product detected by the PB1.7 hybridization probe which is denoted by an open rectangle below the line.

PhosphorImager units of the corresponding 3.3 kb 5'Pst-3'Pst band; thus the formula for calculating percentage of total units in the Msp-3'Pst sub-band is: 100 X [(Msp sub-band) - background] / ([(Msp sub-band) - background] + (Pst band)). This value for 5aCdr-treated samples digested with 200 units of Msp I was used for the graph in Figure 4.6.

RT-PCR of Human HPRT mRNA

Detecting the appearance of human HPRT mRNA was performed by reverse transcriptase-PCR (RT-PCR) as described by Sasaki et al. (133). The position of the forward HPRT RT-PCR primer (TCCTCCTGAGCAGTCAGC) is shown in Figure 4.1. The reverse primer (GGCGATGTCAATAGGACTC) is located in exon 9. First strand cDNA was reverse-transcribed from total RNA with random hexamer primers and amplified by PCR with human HPRT-specific primers. PCR products were fractionated on agarose gels, Southern blotted and hybridized with a radiolabeled human HPRT cDNA probe; the human HPRT-specific RT-PCR product is 794 bp.

Quantitation of hybridization signals in Southern blots was performed by PhosphorImager analysis. The relative levels of human HPRT mRNA in representative samples on the Southern blot in Figure 4.6 were calculated as follows. The radioactivity in each 794 bp HPRT RT-PCR product was quantitated by PhosphorImager analysis and expressed in PhosphorImager units; this was determined for each 5aCdr-treated sample (at 0, 12, 24, 32, 48, and 60 hr) using 0.5 ug of RNA in the RT-PCR reaction and averaging results of duplicate RT-PCR reactions. RNA from a human-rodent somatic cell hybrid containing an active human X chromosome, GM06318, was treated similarly (see Figure 4.3). Total RNA for the GM06318 sample was normalized to that of the experimental samples according to the amount of cell surface antigen 12e7 gene (MIC2)

RT-PCR product. HPRT values for the GM06318 sample were considered to represent 100 percent reactivation. General non-specific background for the autoradiogram was determined by quantitating and averaging the PhosphorImager units at seven separate points on the autoradiogram that did not contain experimental samples; this number (background) was subtracted from the PhosphorImager values obtained for each of the experimental samples to yield "net units." Net units were then divided by the micrograms of RNA template (ug RNA) used in each RT-PCR reaction to yield "(net units)/(ug template)." The values for "percentage of maximal" were then determined by dividing (net units)/(ug template) for each 5aCdr-treated sample by the (net units)/(ug template) obtained from GM06318 control RNA, and multiplying by 100. Thus, the formula used was: Percentage of maximal = [100 X (units 5aCdr - background)/(ug RNA)] / [(units GM06318 - background)/(ug RNA)]

Preparation of DNA: In Vitro Dimethyl Sulfate (DMS) Treatment and DNA Isolation

DMS treatment of cells, DNA purification, and piperidine cleavage of DMS-treated DNA were performed essentially as previously described (62, 63). 5aCdr-treated cells were in vivo footprinted by treatment with 0.2% DMS for 5 minutes in RPMI growth medium.

Ligation-Mediated PCR (LMPCR) and Detection of In Vivo Footprints

LMPCR was carried out essentially as described by Hornstra and Yang (60, 62, 63). For LMPCR, primer set E was used to analyze the upper strand encompassing the -91 footprint, and primer set C was used to analyze the upper strand in the region of the GC boxes, as previously described for in vivo footprinting of the human HPRT gene (63). The positions of the LMPCR primers (and the regions they analyze) and the locations of the footprinted sites used in this study are shown in Figure 4.1. Reaction conditions for first strand synthesis, ligation, and PCR amplification were identical to those previously described (60).

Subsequent gel electrophoresis and electroblotting were carried out as previously described, using a 5% Long Ranger gel (AT Biochem) substituted for the standard polyacrylamide DNA sequencing gel (60, 62, 63). To visualize the final DNA sequencing ladder, strand-specific hybridization probes were synthesized from M13 clones containing the human HPRT promoter region. Probe synthesis, hybridization, washing, and autoradiography were performed as previously described (60, 62, 63).

Quantification of Footprints by Autoradiogram Densitometry

Quantification of transcription factor binding at position -91 was carried out by densitometry of the DNA sequencing autoradiogram. The densitometric value for the band intensity at position -91 in each sample was first normalized for loading differences using the average band intensity of 8 non-footprinted bands flanking both sides of the band at position -91 in each lane. The intensity of the -91 footprint was then expressed as a percentage of the average intensity of the same band in control cell lines 4.12 and GM00468 (Fig. 4.4, lanes 6 and 7), each of which carries an active human X chromosome. The basal band intensity at position -91 in naked DNA (Fig. 4.4, lane 1) was subtracted from each normalized value (lanes 2-5) and averaged control samples (lanes 6 and 7) prior to calculating the final percentage. Quantification of bands at positions -198 and -210 was carried out in the same manner except that the average band intensity of 4 non-footprinted flanking bands was used to normalize for loading differences.

Three other DMS footprints found within the active X and within the 5-azacytidine-reactivated HPRT promoters (63) are also potentially footprinted sites in the 5aCdr-treated X8-6T2 cells. However, these sites are not as strongly footprinted in cells that express HPRT fully (63), and the percentage of reactivated cells at any of the time points examined following 5aCdr treatment is relatively low. Therefore, it could not be determined by LMPCR in vivo footprinting, whether or not any of these weaker footprints had appeared in a fraction of the X8-6T2 population following 5aCdr treatment.

Results

During the course of 5aCdr-mediated reactivation of genes on the inactive human X chromosome, the human HPRT gene was monitored for changes in chromatin structure, appearance of human HPRT mRNA, and binding of transcription factors to the promoter region. A hamster-human somatic cell hybrid containing an inactive human X chromosome (X8-6T2) was first treated with 5aCdr for 24 hr, then returned to normal growth medium without 5aCdr. At various times from 0 to 60 hr after the addition of 5aCdr, chromatin structure of the 5' region was examined by nuclease sensitivity, human HPRT mRNA was detected by RT-PCR, and the binding of transcription factors was assayed by DMS in vivo footprinting. These assays were performed simultaneously on the same cultures of 5aCdr-treated cells.

Changes in Chromatin Structure of the 5' Region During 5aCdr-Induced Reactivation of the HPRT Gene

Discrete regions of nuclease sensitivity in the 5' region of HPRT are known to be present only on the active gene, and these can be detected with either Msp I or DNase I (85, 133, 152, 155). To examine the effect of 5aCdr on the chromatin structure of the

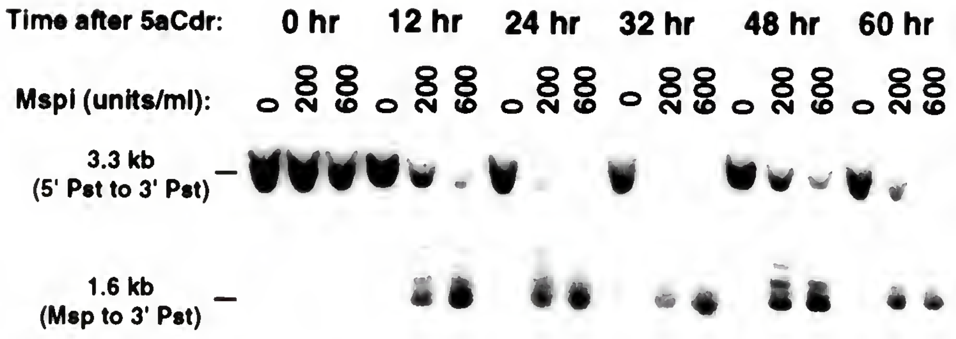


FIGURE 4.2 - Nuclease sensitivity of the 5' CpG island of the human HPRT gene following treatment of X8-6T2 cells with 5aCdr. Nuclei were isolated from X8-6T2 cells 0, 12, 24, 32, 48, and 60 hrs after initiating treatment with 5aCdr. Nuclei from each time point were treated with 0, 200 and 600 units/ml of Msp I. Genomic DNA was isolated, digested to completion with Pst I, and subjected to Southern blot analysis with human HPRT hybridization probe PB1.7 as described by Sasaki et al. (115). The 3.3 kb PstI parental band (denoted *5'Pst to 3'Pst*) contains numerous HpaII/MspI sites in the 5' CpG island of the human HPRT gene. The major 1.6 kb fragment (denoted as *Msp to 3'Pst*) results from cleavage of the parental 3.3 kb PstI fragment within intact nuclei at a cluster of MspI sites in the first exon.

inactive HPRT gene, treated and untreated X8-6T2 nuclei were digested with Msp I as previously described by Sasaki et al. (133). As shown in Figure 4.2, sensitivity of chromatin in the 5' CpG island was negligible in X8-6T2 cells prior to 5aCdr treatment (0 time), became near maximal by 12 hr, and reached maximal sensitivity to digestion with Msp I by 24-32 hr after initiating treatment with 5aCdr. Thus, the major chromatin structure changes in the HPRT gene 5' region, as reflected by the increased accessibility of Msp I sites in chromatin to cleavage by Msp I, occurred within 24-32 hr of exposure to 5aCdr. These results are similar to those reported by Sasaki et al. (133) where maximal sensitivity of chromatin to Msp I digestion was achieved by 24 hr after initiating 5aCdr treatment.

Appearance of Human HPRT mRNA During 5aCdr-Induced Reactivation of the HPRT Gene

The appearance of human HPRT mRNA in the same populations of 5aCdr-treated X8-6T2 cells was assayed by RT-PCR of total RNA. Figure 4.3 shows Southern blot analysis of human HPRT RT-PCR products amplified from total RNA of 5aCdr-treated samples using a radiolabelled human HPRT cDNA hybridization probe. Human HPRT mRNA first became detectable at 24 hr following addition of 5aCdr, and reached maximal levels at 60 hr, when the experiment was terminated. Thus, detectable HPRT mRNA levels did not begin to appear until the chromatin structure of the 5' region had nearly reached its maximal sensitivity to MspI, an observation similar to that of Sasaki et al. (133).

Binding of Transcription Factors During 5aCdr-Induced Reactivation of the HPRT Gene

The binding of transcription factors to the 5' region during reactivation of the inactive HPRT gene by 5aCdr treatment was assayed (in the same populations of 5aCdr-

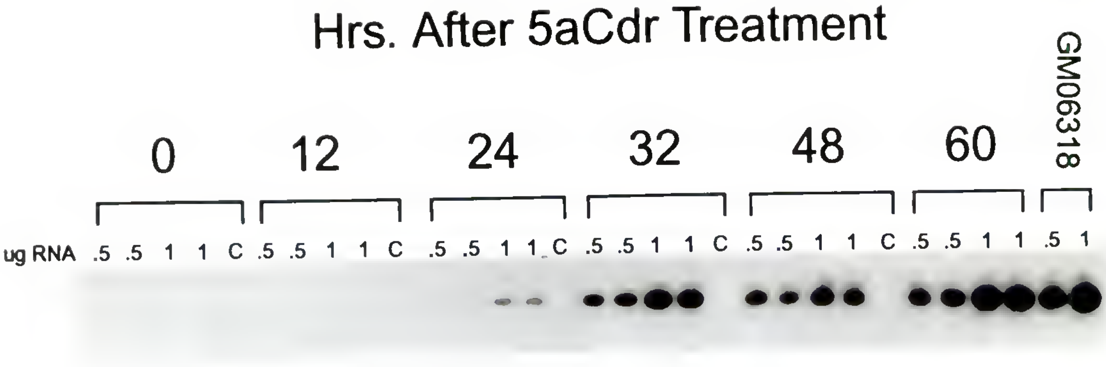

Hrs. After 5aCdr Treatment
0
12
24
32
48
60
GM06318
ug RNA
.5 .5 1 1 C .5 .5 1 1 C .5 .5 1 1 C .5 .5 1 1 C .5 .5 1 1 C .5 .5 1 1 .5 1

FIGURE 4.3 - RT-PCR analysis of human HPRT mRNA in X8-6T2 cells following treatment with 5aCdr. Levels of human HPRT mRNA were assayed by RT-PCR at 0, 12, 24, 32, 48, and 60 hrs after initiating 5aCdr treatment of X8-6T2 cells (lanes marked by numbered brackets). Numbers above each lane denote micrograms of RNA used in RT-PCR reaction; RNA from each 5aCdr-treated sample was assayed by RT-PCR in duplicate. *C* indicates control lanes in which no RNA was added to the RT-PCR reactions. GM06318 is a rodent-human somatic cell hybrid carrying an active human X chromosome.

treated cells as above) by ligation-mediated PCR (LMPCR) in vivo footprinting (60, 62, 63). In previous studies, an in vivo dimethyl sulfate (DMS) footprint in the 5' region of the human HPRT gene was detected at position -91 (relative to the translation start site) on the active human X chromosome (63); this footprint was not detected on the inactive HPRT allele. The identical footprint was also observed in 5aC-treated cells that were HAT (hypoxanthine-, aminopterin-, and thymidine-containing medium)-selected for reactivation of the human HPRT gene on the inactive X chromosome (63). This footprint is characterized by a band of strongly enhanced autoradiographic intensity at position -91 in the guanine-specific DNA sequencing ladder (as compared to the intensity of the same band in the transcriptionally inactive allele or in naked DNA), indicative of a very DMS-reactive guanine residue on the active allele due to the binding of a transcription factor in vivo. The unusual strength of enhanced DMS reactivity at this position on the transcriptionally active HPRT allele allows detection of the bound transcription factor in vivo despite reactivation of the HPRT gene only in a fraction of 5aCdr-treated cells (40).

To examine the binding of transcription factor(s) to the -91 region during reactivation of the human HPRT gene, X8-6T2 cells were assayed by DMS in vivo footprinting at 0, 12, 24, 32, 48, and 60 hrs after addition of 5aCdr in the same samples assayed for MspI sensitivity and HPRT mRNA. Intact 5aCdr-treated cells were treated with DMS to generate in vivo footprints as described previously, and the -91 footprint was visualized by LMPCR using primer set E (63). Given the unusually strong signal exhibited by the footprint at position -91, the binding of transcription factor(s) to this site on the HPRT gene promoter in 5aCdr-treated cells can be detected by an increase in the

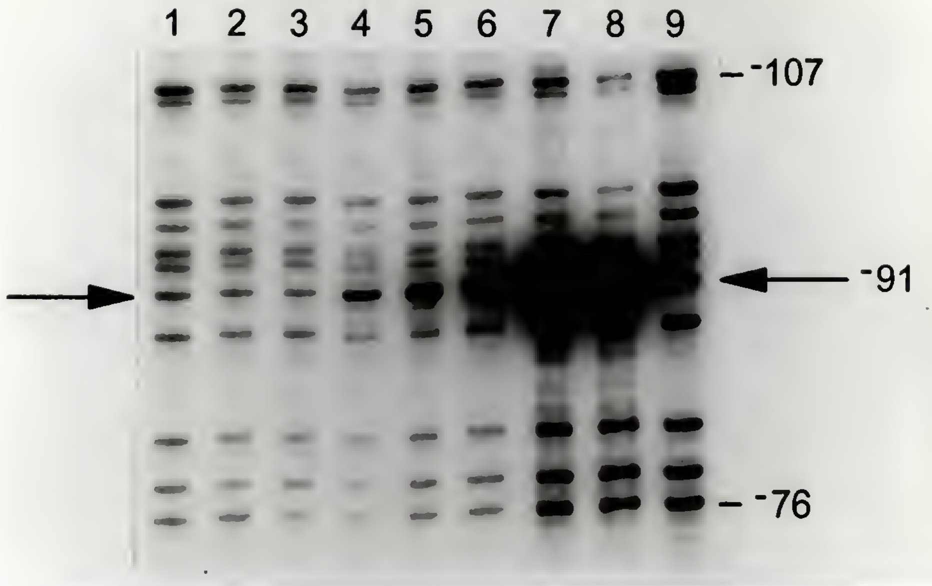


FIGURE 4.4 - In vivo footprint analysis at position -91 following treatment of X8-6T2 cells with 5aCdr. Lane 1, no 5aCdr treatment; lane 2, 12 hr after initiating 5aCdr treatment; lane 3, 24 hr after initiating 5aCdr treatment; lane 4, 32 hr after initiating 5aCdr treatment; lane 5, 48 hr after initiating 5aCdr treatment; lane 6, 60 hr after initiating 5aCdr treatment; lane 7, GM00468 cells; lane 8, 4.12 cells, lane 9, 8121 cells. Arrows indicate the position of the highly DMS-reactive guanine indicative of the -91 footprint, and numbers indicate nucleotide position relative to the translation start site. GM00468 and 4.12 cells are positive controls containing an active human X chromosome, and 8121 cells are a control containing an inactive human X chromosome.

relative intensity of the band at position -91 in the final guanine-specific DNA sequencing ladder.

Figure 4.4 shows the results of in vivo footprinting assays on the 5aCdr-treated cells in the -91 region. Lanes 7 and 8 show the very intense DMS modification and cleavage of the guanine residue at position -91 that is indicative of the in vivo footprint in two control cell cultures containing only an active human X chromosome. Lane 1 displays the DMS modification and cleavage pattern of the HPRT 5' region in X8-6T2 cells (containing an inactive human X chromosome) prior to 5aCdr treatment. This sequencing ladder is typical of the in vivo guanine-specific modification and cleavage pattern seen for the transcriptionally inactive human HPRT gene (63) as well as for naked genomic DNA purified prior to DMS treatment. All samples assayed from 0-24 hr (lanes 1-3) after exposure to 5aCdr produced the guanine-specific cleavage pattern of the control sample, with the band intensity at position -91 identical to that of naked DNA (data not shown). Thus, no in vivo footprint at position -91 is detectable up to 24 hr after addition of 5aCdr. However, beginning at 32 hr (lane 4) and continuing through 60 hr (lane 6) following addition of 5aCdr, a gradual increase in the intensity of the band corresponding to the -91 footprint is detected relative to the intensity of adjacent and surrounding non-footprinted bands in the sequencing ladder (lanes 4-6). Densitometric analysis of relative band intensity at position -91 is shown in Figure 4.6. Thus, the binding of a transcription factor(s) at position -91 appears to correlate most closely with the appearance of HPRT mRNA, reaching its maximum level at 60 hr, rather than with the alteration of chromatin structure which reaches a maximum at 24-32 hr. Thus, the remodeling of chromatin structure of the 5' region to a more nuclease-sensitive conformation in response to 5aCdr treatment does not require binding of the transcription factor(s) to the region of the promoter surrounding position -91.

A similar result is also seen upstream of position -91 in a region containing 5 GC boxes, potential binding sites for the transcription factor Sp1 (11). This region exhibits multiple footprints on the active HPRT allele that include three guanines with enhanced DMS reactivity on the upper strand at positions -163, -198, and -210 (63). These sites of enhanced DMS reactivity are not detected on the inactive HPRT allele or on naked DNA purified prior to DMS treatment. We chose to analyze sites of enhanced DMS reactivity because they are likely to be more readily detectable in the subpopulation of 5aCdr-treated cells that reactivate the HPRT gene than sites that exhibit protection from DMS reactivity. Figure 4.5 shows the results of LMPCR in vivo footprinting of the region containing positions -198 and -210 in X8-6T2 cells following 5aCdr treatment (using LMPCR primer set C; (63)). As with the footprint at position -91, the enhanced site of DMS reactivity at position -198 shows a detectable increase in intensity at 60 hr. (lane 6) after addition of 5aCdr as compared to the sample not treated with 5aCdr (lane 1) and time points prior to 32 hr (lanes 2 and 3). Because the enhanced DMS reactivity at position -198 on the active HPRT allele is weaker than the enhanced reactivity seen at position -91, the increase in band intensity with time at position -198 is not as strong as the increase in band intensity at position -91 in Figure 4.4. Nonetheless, by comparing the relative intensity of the band at position -198 to the intensity of the adjacent band at position -199, it is clear that by 60 hr (lane 6) the relative intensity of the band at position -198 is stronger than that at position -199. This relative increase in band intensity at position -198 compared to position -199 is consistent and reproducible. The relative

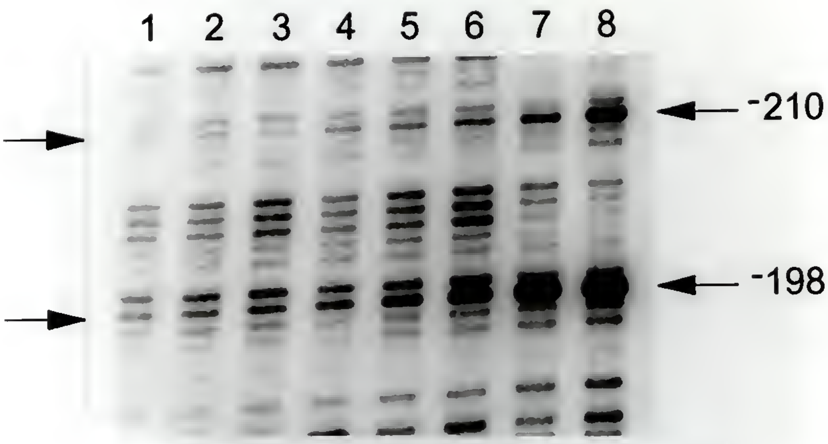


FIGURE 4.5 - In vivo footprint analysis of the GC box region following treatment of X8-6T2 cells with 5aCdr. Autoradiogram of in vivo footprints at positions -198 and -210. Lane 1, no 5aCdr treatment; lane 2, 12 hr after initiating 5aCdr treatment; lane 3, 24 hr after initiating 5aCdr treatment; lane 4, 32 hr after initiating 5aCdr treatment; lane 5, 48 hr after initiating 5aCdr treatment; lane 6, 60 hr after initiating 5aCdr treatment; lane 7, 4.12 cells; lane 8, GM00468 cells. Arrows indicate positions of enhanced in vivo DMS reactivity (in vivo footprints) on the active HPRT allele, and numbers indicate nucleotide position relative to the translation start site.

difference in band intensity at position -198 compared to the band intensity at position -199 is not evident at time points prior to 32 hr (lanes 1-3). Therefore, as seen with the -91 footprint, these data also indicate that the binding of a transcription factor at position -198 (most likely Sp1) occurs late in the process of 5aCdr-mediated reactivation (well after maximal levels of nuclease sensitivity have been achieved at 24-32 hr) and correlates more closely in time with active transcription of the HPRT gene and binding of factors at position -91 rather than alteration in the chromatin structure of the HPRT locus.

The site of enhanced DMS reactivity at position -210 in the GC box region also exhibits an increase in intensity (relative to the adjacent band at position -211) at both 48 hr (lane 5) and 60 hr (lane 6) after 5aCdr treatment, while no such relative increase is detected prior to 32 hr (lanes 1-3). Again, this would indicate the binding of a transcription factor at this position late in the 5aCdr reactivation process after chromatin remodeling has occurred. Densitometric analysis of relative band intensities at positions -198 and -210 during 5aCdr reactivation is shown in Figure 4.6.

In contrast, the site of enhanced DMS reactivity at position -163 in the GC box region does not exhibit a clear increase in intensity (relative to the adjacent band at position -164) during the course of 5aCdr reactivation (data not shown). The inability to demonstrate clear evidence of this footprint during 5aCdr reactivation most likely reflects limitations on the sensitivity of the in vivo footprinting assay at this site (as discussed in Materials and Methods).

A graphical summary of the events we studied following 5aCdr treatment of the inactive X hybrid is shown in Figure 4.6, in which chromatin structure (nuclease sensitivity), transcription factor binding at positions -91, -198, and -210, and HPRT

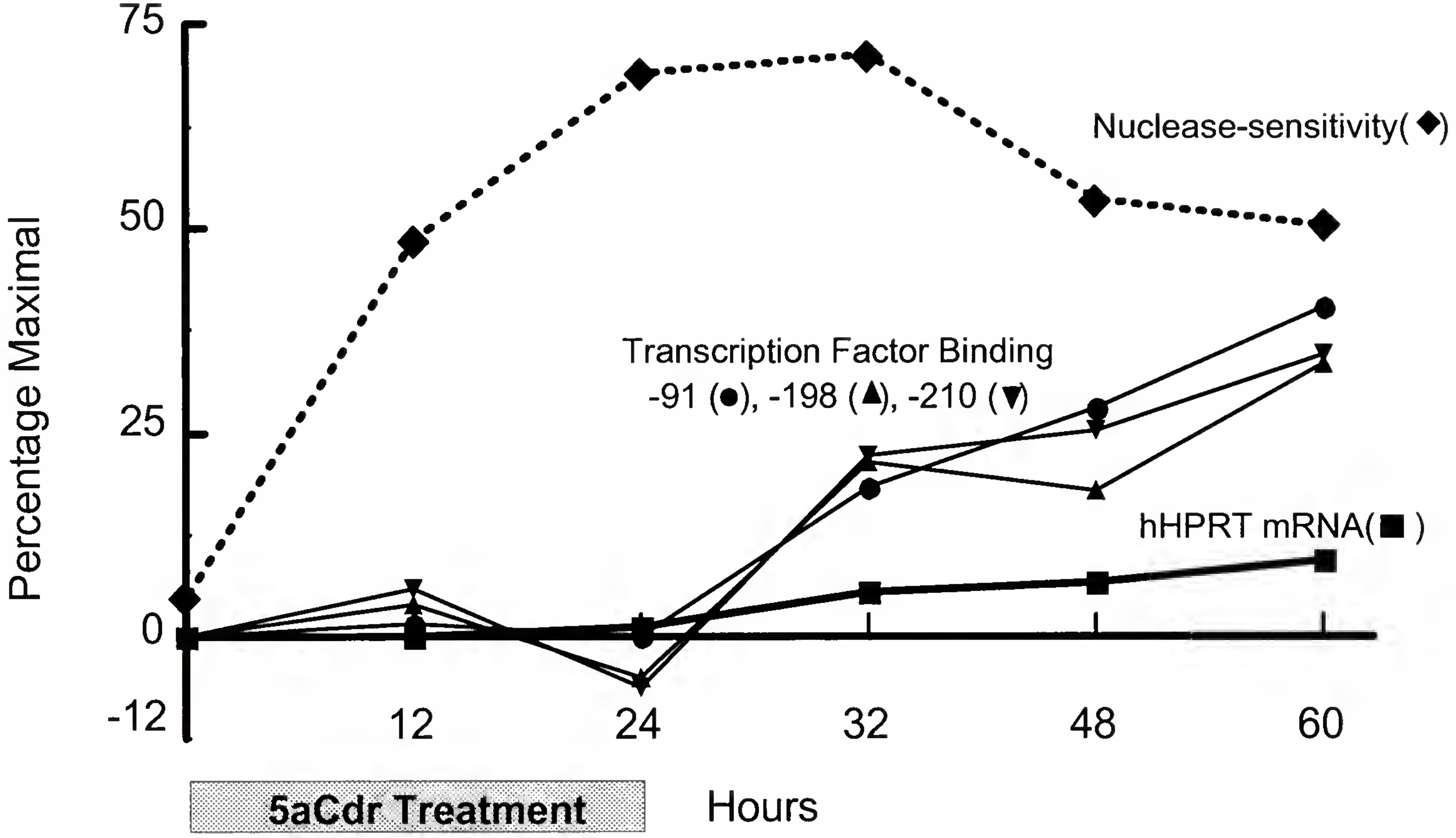


FIGURE 4.6 - Time course of HPRT mRNA levels, nuclease sensitivity of chromatin, and transcription factor binding at positions -91, -198, & -210 following 5aCdr treatment. Nuclease-sensitivity denotes the percentage of the 3.3 kb PstI fragment digested by MspI within intact nuclei; calculations for the graph are based on samples digested with 200 units of MspI in Figure 4. 2. hHPRT mRNA indicates the percentage of human HPRT mRNA in 5aCdr-treated samples relative to human HPRT mRNA detected in cultured cells containing an active X

mRNA levels are plotted as a function of time after initiating 5aCdr treatment. The appearance of the -91, -198, and -210 footprints are correlated with the appearance of HPRT mRNA rather than with the earlier change in chromatin structure. This change in chromatin structure, therefore, does not require binding of a factor(s) to -198, -210 or the -91 region, a region that is near the multiple sites of transcription initiation and whose location is similar to regions previously reported to be critical for silencing other genes by DNA methylation (82, 120). Our data demonstrate that chromatin remodeling in response to 5aCdr treatment does not require transcription factor binding to multiple transcription factor binding sites in the HPRT promoter region.

Discussion

The 5aCdr-induced reactivation of the inactive X-linked HPRT gene involves an initial hemi-demethylation of the promoter region that is associated with a change in chromatin from a nuclease-resistant to a nuclease-sensitive structure. Following the alteration in chromatin structure, symmetrical demethylation occurs and HPRT mRNA appears (133). We show here that transcription factor binding to at least three sites in the promoter region is correlated with the appearance of HPRT mRNA rather than with the preceding remodeling of chromatin structure.

From this sequence of events, we propose that the change in chromatin structure of the 5' region as a result of 5aCdr treatment does not require transcription factor binding in the immediate promoter region (footprints associated with the -91, -198, and -210 sites). Reactivation of HPRT apparently requires a 5aCdr-induced remodeling of chromatin structure such that DNA binding sites in the promoter region become accessible to transcriptional activators. The binding of these activators, which are known to be present in the nucleus prior to 5aCdr treatment (and are bound to the active HPRT allele in female cells), would then affect further changes in chromatin structure of the promoter region that potentiate transcriptional activity (e.g., an alteration in the nucleosomal structure). The primary mechanism by which DNA methylation maintains the silence of the HPRT gene on the inactive X chromosome may, therefore, involve a role in organizing or stabilizing chromatin into a conformation that prevents the accessibility of transcriptional activators, or otherwise precludes their binding to DNA.

Although we were only able to analyze three of the seven sequence-specific footprints characteristic of the HPRT gene on the active X chromosome (63), the recent report of Martinez-Balbas et al. (92) that nuclease hypersensitivity of the human hsp70 gene is independent of transcription factor binding suggests that if it were possible to analyze the remaining footprints, these would also appear after the induction of chromatin remodeling by 5aCdr treatment. It is possible that DNA-binding proteins not detected by our earlier DMS in vivo footprinting studies of the promoter region of the active and inactive HPRT alleles could be responsible for remodeling of chromatin structure in response to 5aCdr. However, recent DNase I in vivo footprinting of the human HPRT gene on the active and inactive X chromosome shows no evidence for sequence-specific DNA-protein interactions on the upper strand between positions -77 and -227 on the inactive allele, nor does it show additional footprints in regions not previously footprinted by DMS on the active allele [(63); and C. Chen, Y. Yu, and T.P. Yang, in preparation]. These DNase I studies are consistent with our original findings that only the GC boxes, AP2 site, and the transcription initiation region are bound by factors on the active HPRT allele, and that no DNA-binding proteins are found on the inactive allele in this region (including methyl DNA-binding proteins). The same conclusions were reached by Pfeifer et al. (119, 122) in similar in vivo footprinting studies of the human PGK-1 gene using DMS and DNase I.

It has been postulated that DNA methylation affects transcription by directly altering the interaction of DNA-binding regulatory proteins with their binding sites (39), or by altering chromatin structure and secondarily altering sequence-specific DNA-protein interactions (19, 73). Recent studies of high resolution methylation patterns in the human HPRT and the human and mouse PGK-1 gene 5' regions suggest that direct, methylation-induced alteration of sequence-specific DNA-protein interactions in the promoter region is unlikely to be the primary mechanism by which X-linked genes are silenced by DNA methylation. In these studies methylation patterns in the 5' region of the human and mouse PGK-1 genes (120, 122, 143) and human and mouse HPRT genes (63, 87) reveal no strict correlation between methylated CpG dinucleotides and binding sites for transcriptional activators, and no discernable conserved pattern of CpG methylation among the inactive alleles of these genes (other than a generally higher level of methylation compared to the active allele). In particular, the methylation pattern seen in the mouse PGK-1 gene makes it unlikely that DNA methylation functions by directly modifying individual interactions between transcriptional activators and their DNA binding sites in the immediate promoter since only a single CpG dinucleotide is fully methylated on the inactive X chromosome (143). The results we report here support the concept that DNA methylation and demethylation primarily affect chromatin structure and secondarily (as a result of changes in chromatin structure) influence sequence-specific DNA-protein interactions of the promoter region within native chromatin.

CHAPTER 5
TRANSCRIPTIONAL INACTIVATION OF THE MOUSE HPRT GENE DURING SPERMATOGENESIS

Introduction

Most male cells are not subject to X chromosome inactivation (46). However, the cells of the male germ line appear to undergo X chromosome inactivation during spermatogenesis. This event is distinct from the genome-wide transcriptional silencing of genes by protamine replacement of histones which occurs at a later stage (75). While RT-PCR analysis of five X-linked genes (97, 136) suggest that the male X chromosome is transcriptionally silent at the round spermatid stage of spermatogenesis, the exact timing of this inactivation is unclear due to the stability of mRNA.

Several lines of evidence indicate that the binding of sequence-specific DNA binding proteins is correlated with gene expression (91, 105). Consequently, the transcriptionally activity of a gene can be determined by monitoring transcription factor binding to its promoter. Since the DMS in vivo footprint pattern of the mouse HPRT gene is known (87), this gene was chosen to monitor the transcriptional status of the male X chromosome during spermatogenesis. Three of the four footprinted regions on active mouse HPRT gene were examined (87). Two of these footprinted regions occur over GC boxes and contain 4 guanine residues that show enhanced DMS reactivity. Another footprint, which also shows enhanced DMS reactivity, occurs near the transcription initiation start site at +14. No footprints were observed on the inactive X chromosome. Therefore, at these sites, an active mouse HPRT gene will have transcription factors bound to it and an inactive gene will have no factors bound. By following the appearance and disappearance of these 5 footprinted guanine residues, the degree of binding of transcription factors to the HPRT gene can be monitored in spermatogenic cells. Thus, the time at which the mouse HPRT gene becomes transcriptionally inactivated can be examined. In order to determine the precise stage(s) at which X inactivation occurs during spermatogenesis, the HPRT gene in primitive type A spermatogonia, preleptotene spermatocytes, leptotene+zygotene spermatocytes, early pachytene spermatocytes, late pachytene spermatocytes, round spermatids, testicular spermatozoa and caput epididymal spermatozoa from the mouse was examined by ligation-mediated PCR (LMPCR)-assisted DMS in vivo footprinting.

Materials and Methods

Separation of Spermatogenic Cells

Male mice of the appropriate ages (6-60 days) were sacrificed and the testes were excised, decapsulated and dispersed as described by Bellve et al. (129) and Romrell et al. (4). Dissociated seminiferous epithelium were separated by sedimentation velocity at unit gravity at 4°C, using a 2-4% BSA gradient in EKRB (120.1 mM NaCl, 4.8 mM KCl, 25.2 mM $NaHCO_3$, 1.2 mM KH_2PO_4, 1.2 mM $MgSO_4 \cdot 7H_2O$, 1.3 mM $CaCl_2$, 11.1 mM glucose, 1 mM glutamine, 10 ml/liter of essential amino acids, 10 ml/liter of nonessential amino acids, 100 ug/ml streptomycin sulfate, and 60 ug/ml penicillin G (K^+ salt). Cell types were identified on the basis of morphology using phase contrast microscopy. For spermatozoa, the heads were prepared as sonication-resistant cells from which nuclei were prepared. Spermatozoa, round spermatids and adult pachytene spermatocytes were greater than 90% pure. All other purified cell populations were greater than 85% pure.

DMS Treatment

Purified cells (approximately $5x10^6$) were incubated in 0.5% DMS for 90 seconds at room temperature. The reaction was stopped by adding 49 ml of ice-cold PBS, pelleting the cells, and washing in 50 ml of ice-cold PBS. Cells were recovered by centrifugation, and lysed in 5 ml DNA extraction buffer (50 mM Tris-HCl pH 8.5, 50 mM NaCl, 25 mM EDTA, 0.5% SDS and 300 ug/ml proteinase K) at 37°C overnight. DNA was purified by successive phenol and chloroform extractions. Following extraction, the purified DNA was precipitated with 2 volumes of 100% ethanol and 0.2 M NaCl, washed with 80% ethanol, resuspended in 1M piperidine, and incubated at 95°C for 30 minutes. Cleaved DNA was dried in a vacuum concentrator and resuspended in 1xTE to a concentration of 1ug/ul.

Ligation-Mediated PCR

LMPCR was carried out as described by Litt et al. (87). For LMPCR, the lower strand was analyzed with primer set J and the upper strand was analyzed with primer set U (see materials and methods, Chapter 2). Reaction conditions for first strand synthesis and ligation were identical to those previously described (87). PCR condition were the same as previously described (87). The gel electrophoresis, electrotransfer, hybridization, and probing was carried out as described by Hornstra and Yang (63).

Quantification of Footprints by Autoradiogram Densitometry

Quantification of transcription factor binding at positions +14, -38, -39, -52, and -51 was carried out by densitometry of the DNA sequencing autoradiogram. The densitometric value for the band intensity at each position in each sample was first normalized for loading differences using the average band intensity of several non-footprinted bands flanking both sides of the band at each position in each lane. The intensity of the footprint was then expressed as a percentage of the average intensity of the same band in control cell samples relative to the in vitro and in vivo treated spleen samples. The in vitro treated spleen represented 0% footprinted levels and in vivo treated spleen, which contains an active HPRT gene, represented 100% footprinted levels. The basal band intensity at each position in naked DNA was subtracted from each normalized value and from the control samples prior to calculating the final percentage.

Results

The binding of transcription factors to the promoter of genes has been shown to correlate with gene expression (91, 105). Previously, in vivo footprinting of the active and inactive mouse HPRT gene (87) in female cells revealed several footprints on the active allele and no footprints on the inactive allele. Thus, it is likely that the absence of any detectable footprints in the 5' region of the mouse HPRT gene on the male X chromosome would indicate that the gene has been transcriptionally inactivated. Previous studies indicate that several gene are undergoing transcriptional silencing at approximately the same stage of spermatogenesis (97, 136). Therefore, it is probable that the transcriptional silencing of one X-linked gene may reflect the transcriptional silencing of other X-linked genes. In order to determine the precise stage at which genes on the male X chromosome become inactivated during spermatogenesis, we have employed LMPCR-assisted DMS in vivo footprinting to examine the binding of transcription factors within the 5' region of the mouse HPRT gene during different stages of spermatogenesis. Five different footprinted guanine residues that showed enhanced DMS modification were used as indicators of transcription factor binding. Primer set U was used to examine the footprints in the GC boxes on the upper strand and primer set J was used to examine the footprint at +14 on the lower strand (Figure 5.1).

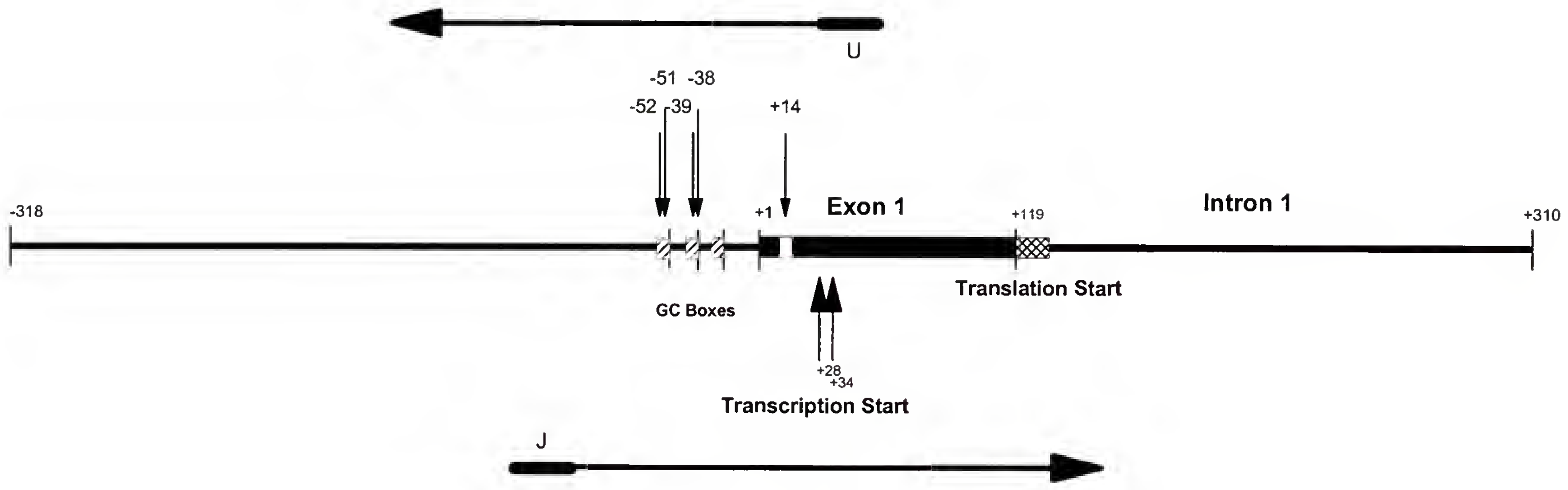
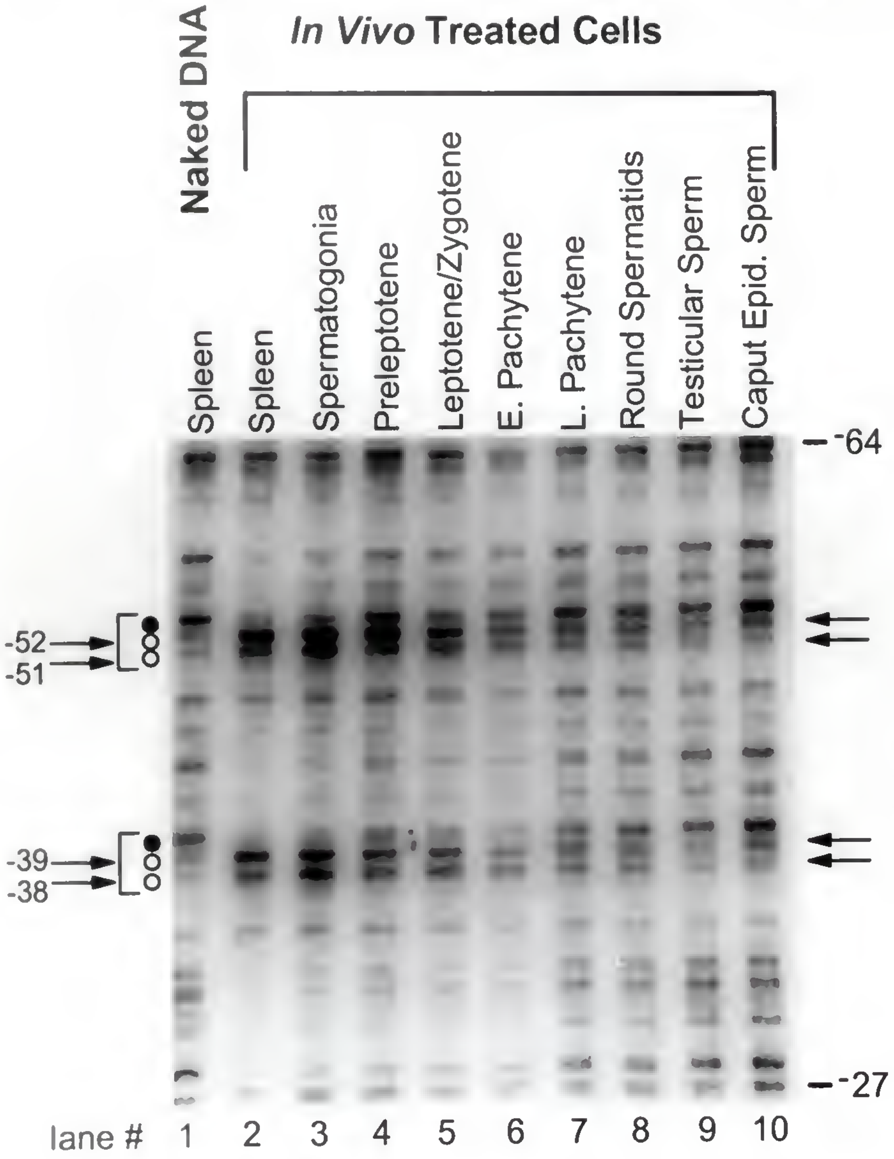


FIGURE 5.1 - Location of LMPCR primers for analysis of the mouse HPRT 5' region. The numbered horizontal line represents the mouse HPRT gene 5' region with positions numbered relative to the major transcription start site. The large rectangle represents the first exon with the solid portion signifying the 5' untranslated region. The smaller open boxes represent GC boxes (GGGCGG). The open box indicates the novel 9-bp sequence which is identical in both the human and mouse HPRT genes. The crosshatched box indicates the translated portion of the first exon. The Number vertical arrow designated the position of the 5 footprint examined in these experiments. The short heavy lines above and below the numbered line indicate the position of the U and J PCR primer used in the LMPCR footprinting analysis. The large arrows indicate the region and direction resolved by each primer set.

Transcription factor binding to the upper strand of the mouse HPRT gene is examined at different stages of spermatogenesis in Figure 5.2. The first lane designated "Naked DNA (spleen)" contains DNA that was first purified from male mouse spleens and then treated in vitro with DMS. Therefore, this sample shows a pattern of DMS reactivity that is not modified by factors interacting with the DNA. This in vitro pattern is identical to that observed in previous in vivo footprinting studies of the mouse HPRT gene (87) on the active and inactive X chromosome (Figure 2.2). The second lane designated "spleen" in the in vivo treated cells shows the footprint pattern of an active HPRT gene. The 4 footprinted guanines designated by open circles are clearly visible. The same footprint pattern is observed on the active HPRT allele in female somatic cells (Figure 2.2).

Examination of these footprinted guanine residues (-52, -51, -39, and -38) in spermatogonia, preleptotene spermatocytes, and leptotene+zygotene spermatocytes suggests that transcription factors are interacting with the HPRT gene and that the HPRT gene is active at these stages. It should be note that the -39 and -52 footprinted guanines do appear to decrease slightly in the intensity of the footprint in early pachytene spermatocytes which suggests that in a small number of the cells the HPRT gene is becoming transcriptionally silent at this stage. The greatest decrease in the intensity of the footprints at -52, -51, -39, and -38 is observed between early pachytene spermatocytes and late pachytene spermatocytes. This suggests the majority of the transcriptional silencing of the HPRT gene is occurring between these stages of spermatogenesis. Except for the -38 footprint, which actually increases in intensity in round spermatids, these footprints (-51, -52, and -39) appear to decrease a slightly in intensity in round spermatids, suggesting that some transcriptional silencing is continuing at this stage. In

FIGURE 5.2 - DMS in vivo footprinting of the upper strand of the mouse HPRT gene at different stages of spermatogenesis. The autoradiograph shows the DMS guanine sequencing ladder of the upper strand of the mouse HPRT gene using primer U. The brackets to the left denote regions of interest. Closed circles indicate footprinted guanine residues which are protected. Open circles indicate footprinted guanine residues which are enhanced. Numbers to the right indicate the position relative to the major transcription start site. Lane designation: NAKED DNA SPLEEN denotes DNA purified from male mouse spleen before in vitro DMS treatment; IN VIVO TREATED CELLS denotes samples in which intact cells are treated in vivo with DMS. SPLEEN-male mouse spleen cells; SPERMATOGONIA-purified primative type A spermatogonia; PRELEPTOTENE-purified preleptotene spermatocytes; LEPTOTENE/ZYGOTENE-purified leptotene+zygotene spermatocytes; E. PACHYTENE-purified early pachytene spermatocytes; L. PACHYTENE-purified adult pachytene spermatocytes; ROUND SPERMATIDS-purified round spermatids; TESTICULAR SPERM-purified testicular sperm; CAPUT EPID. SPERM-purified caput epididymal sperm. All cells were treated with DMS for 90 seconds.


Naked DNA
In Vivo Treated Cells
Spleen
Spleen
Spermatogonia
Preleptotene
Leptotene/Zygotene
E. Pachytene
L. Pachytene
Round Spermatids
Testicular Sperm
Caput Epid. Sperm
-64
-52
-51
-39
-38
-27
lane # 1 2 3 4 5 6 7 8 9 10

testicular sperm and caput epididymal sperm, the footprints at -52, -51, -39, and -38 are undetectable, indicating that the HPRT gene is inactive at these stages. The decrease in intensity of the 4 footprinted bands can be followed more clearly in Figure 5.3, which shows the densitometry of the footprints at different stages of spermatogenesis.

Figure 5.4 examines the binding of transcription factors to the lower strand of the mouse HPRT gene at different stages of spermatogenesis. In previous footprinting analysis of the mouse HPRT gene, a guanine nucleotide showed enhanced DMS reactivity at +14 in samples that contained an active 5' region (Figure 2.3). This same footprint, which is detected as an increase in intensity (darkness) of the guanine band at +14 relative to the intensity of the guanine band directly below it, can be observed in the spleen lane (Figure 5.4, lane 2) of the in vivo treated cells. The guanine residue at +14 is not footprinted in the "Naked DNA spleen" lane (Figure 5.4, lane 1) where the guanine band at +14 is lighter than the guanine band directly below it. The footprint at +14 is detectable in spermatogonia, preleptone spermatocyes, leptotene/zygotene spermatocyes, and early pachytene spermatocytes (lanes 3-6) but undetectable in late pachytene spermatocytes, round spermatids, testicular spermatozoa, and caput spermatozoa (lanes 7-10). Densitometry of the +14 footprint at different stages of spermatogenesis (Figure 5.5) shows a decrease in the intensity of this footprint in for each successive stage of spermatogenesis with the largest change in intensity occurring between spermatogonia and preleptotene spermatocytes. This argues that this site is becoming less accessible to transcription factors in each successive stage of spermatogenesis and that transcriptional silencing of the mouse HPRT gene is occurring thoughout spermatogenesis.

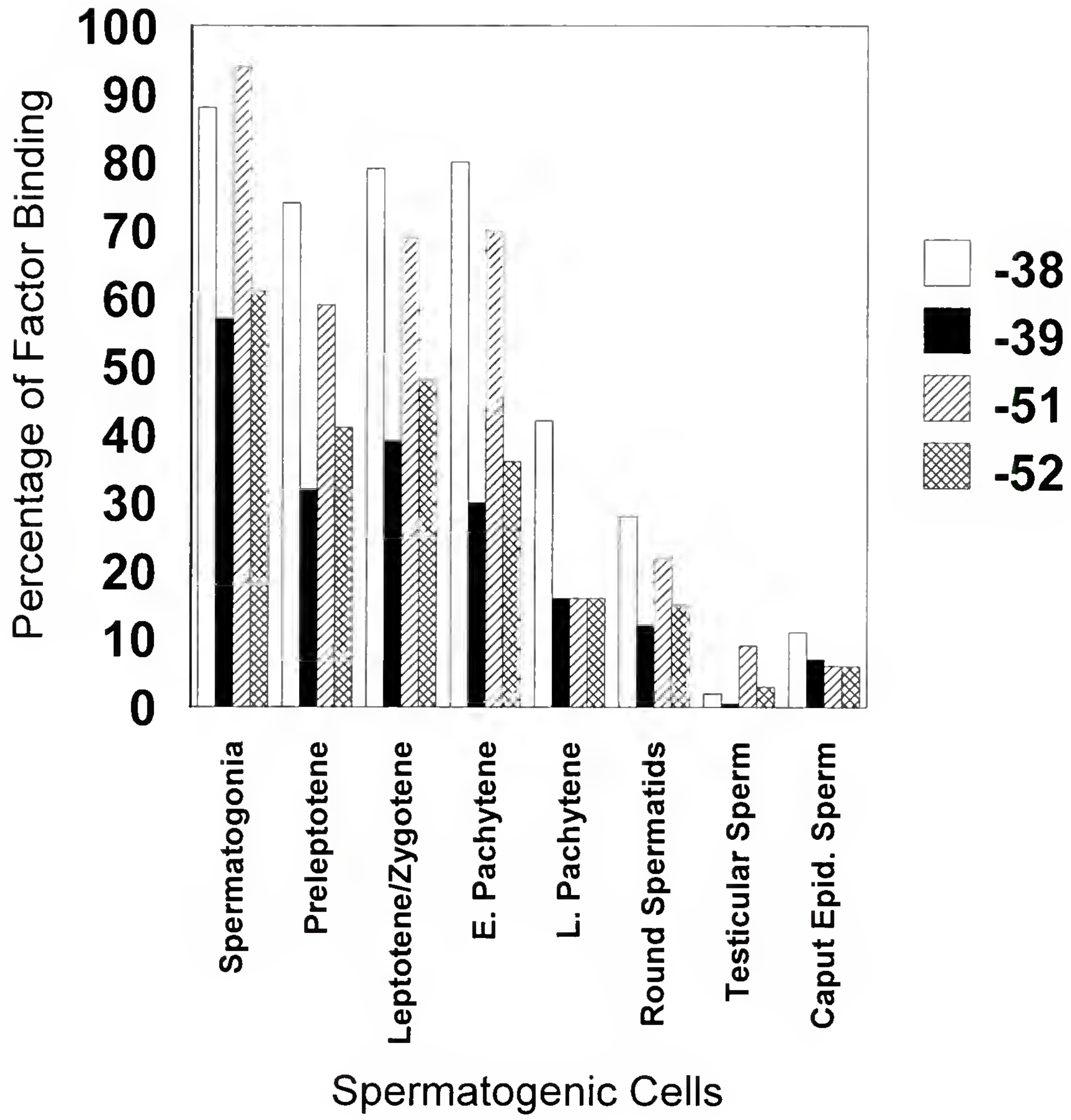


FIGURE 5.3 - Densitometry of four footprints on the upper strand at different stages of spermatogenesis. The histograms shows the relative intensity of four footprinted guanine residues on upper strand the at different stages of spermatogenesis. The legend denotes which histograms represent which guanine. The guanine at -38 is represented by the open histograms, the guanine at -39 is represent by the solid histograms, the guanine at -51 is represented by the diagonally striped histogram, and the guanine at -52 is represent by the cross-hatched histograms. The Y axis indicates the percent of transcription factor binding with the in vivo treated spleen sample being a 100% factor binding and the naked DNA spleen sample being 0% factor binding. The X-axis denotes the different stages of spermatogenesis designation of the stages is the same as Figure 5.2.

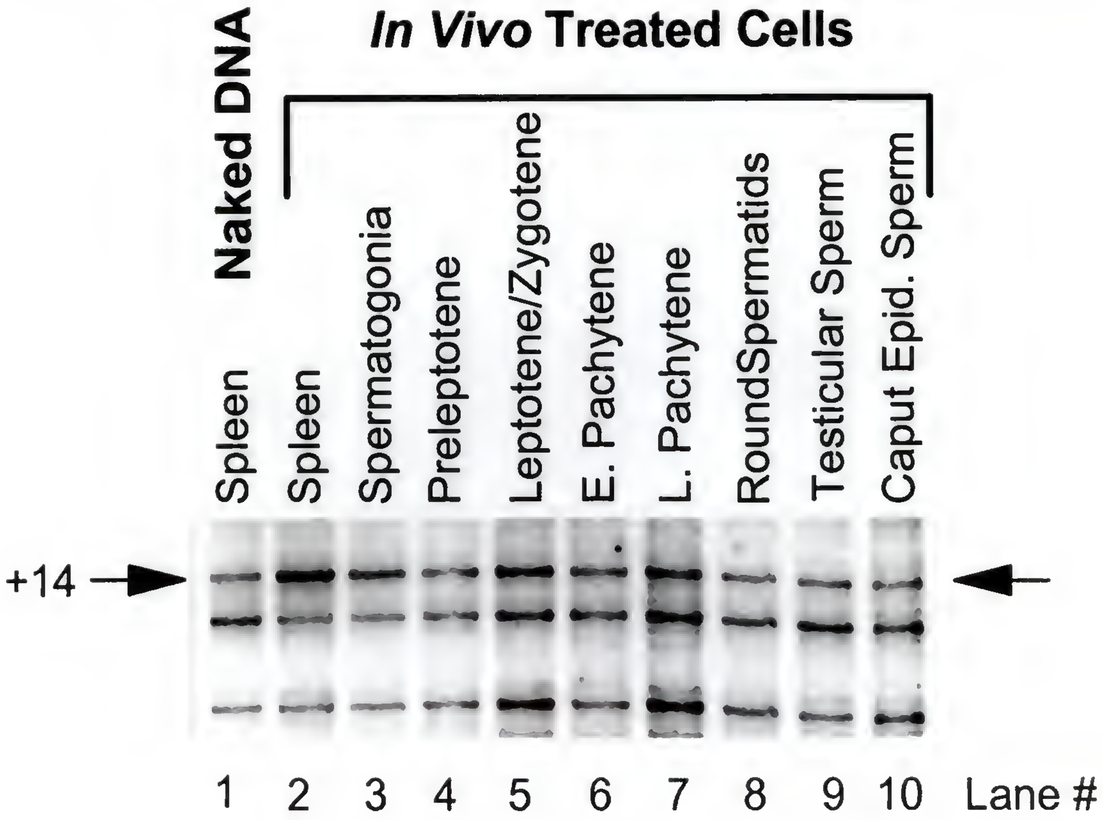


FIGURE 5.4. DMS in vivo footprinting of the lower strand of the mouse HPRT gene at different stages of spermatogenesis. The autoradiograph shows the DMS guanine sequencing ladder of the lower strand of the mouse HPRT gene using primer J. The lane designations and symbols are the same as Figure 5.2.

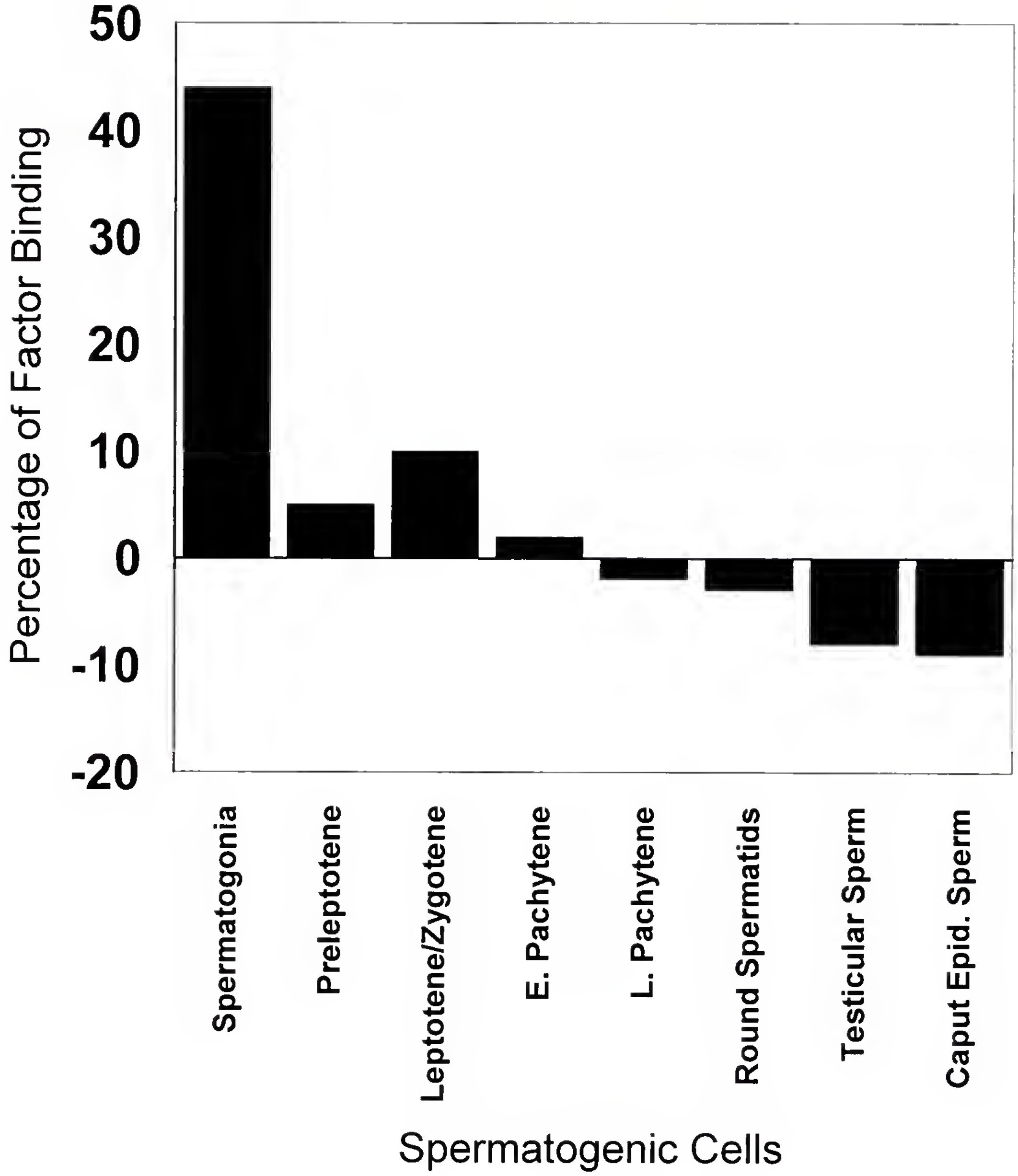


FIGURE 5.5 - Densitometry of the +14 footprint on the lower strand at different stages of spermatogenesis. The histogram shows the relative intensity of the +14 guanine residue at different stages of spermatogenesis. All other designations are the same as Figure 5.3.

Discussion

Examination of the binding of transcription factors to the mouse HPRT gene at different stages of spermatogenesis with DMS in vivo footprinting suggests that the greatest decrease in factor binding occurs between early pachytene spermatocytes and late pachytene spermatocytes. Since transcription factor binding has been correlated with gene expression (91, 105), most of the transcriptional silencing of the mouse HPRT gene is probably occurring between the early and late pachytene spermatocyte stages. However, there appears to be a gradual decrease in the binding of transcription factors to the HPRT gene at each sequential stage of spermatogenesis examined. This suggests that some transcriptional repression of the HPRT gene is occurring throughout the stages of spermatogenesis. A significant decrease in mRNA transcript of the pyruvate dehydrogenase (Pdha-1), phosphorylase kinase (Phka), Pgk-1, zinc finger x (Zfx), and Hprt genes in spermatogenic cells was observed after late pachytene spermatocytes (97, 136). Combined, these data suggests that the male X chromosome is becoming transcriptionally inactive just prior to or during late pachytene spermatogenesis.

Previous in vivo footprinting studies of the mouse HPRT gene on the active and inactive X chromosome in female somatic cells showed four footprinted regions on the active allele and no footprints on the inactive allele (87). The same footprints are observed in spermatogenic cells, although the intensity of the footprints change across the stages. With respect to the types of factors binding at the HPRT promoter, there appears to be no difference between an active male and female HPRT allele or an inactive male and female HPRT allele.

These experiments help verify the time of X inactivation in spermatogenic cells and indicate X inactivation in spermatogenesis leads to a decline in transcription factor binding in meiotic and postmeiotic cells. To further confirm these results nuclear run-off and steady-state cytoplasmic transcripts need to be examined.

CHAPTER 6

IN VIVO FOOTPRINTING OF THE MINIMAL HUMAN XIST PROMOTER ON THE ACTIVE AND INACTIVE X CHROMOSOMES

Introduction

In female mammals, the majority of the genes on one of the two X chromosomes are transcriptionally silenced early in embryogenesis in a process termed X chromosome inactivation (46, 90). This results in dosage compensation of X-linked genes between males and females (46). Previous studies of certain X:autosome translocations indicated that a region, termed the X-inactivation center (XIC), is necessary for X inactivation (16). To date three genes, XIST, CDX4, and BPX, have been mapped within the human and mouse Xic/XIC (14, 58, 130). Of these genes only the X-inactive specific transcript gene (XIST gene) appears to have unique characteristics which suggest it plays an essential role in X inactivation. The XIST gene is expressed only from the inactive X chromosome (14), its RNA transcript is localized to the nucleus and paints only the inactive X chromosome in fluorescence in situ hybridization studies (13, 15, 32), and its expression precedes X inactivation (69). In embryonic stem cells with XIST knock-outs, only the X chromosome containing a functional XIST gene undergoes inactivation (118). Furthermore, a 450 kb murine transgene containing the Xist gene displayed properties of the Xic, including ectopic expression and cis-localization of the Xist transcript, chromosomal counting, and transcriptional silencing of several genes on the recombinant autosome (80). Although several studies have investigated the role of Xist/XIST in X inactivation, little is known about the moleclar mechanisms that regulate XIST expression.

Recent in vitro studies by Hendrich et al. identified multiple regulatory elements in the 5' region essential for XIST expression (56). Deletion and transient expression analyses using a luciferase reporter gene determined that the minimal XIST promoter is located between positions -93 and +31. Saturation site-directed mutagenesis (linker scanning) indicated that three regions are important for normal XIST expression: 1) -78 to -67, 2) -31 to -26, and 3) -1 to +5. Gel mobility-shift studies and nucleotide sequence analyses suggest that the region from -78 to -67 interacts with Sp1 (11), the region from -31 to -26 interacts with TBP (51), and the region from -1 to +5 interacts with YY1 (134). Further mutational analysis suggested that at least two additional sites were important for XIST expression. A significant reduction in XIST expression was generated by in vitro mutagenesis in a second Sp1 site at positions -50 and -51, and within a string of seven cytosines at position -43 (56).

Recently, a rare C to G transversion at position -43 of the XIST promoter has been identified in two independent families which exhibit a skewed pattern of inactivation (109). In female members of this family, the X chromosome carrying this mutation was selectively inactivated, despite the fact that this mutation decreases promoter activity by 5-fold in transient expression assays (56). This suggests that the site containing this mutation may be involved in selecting which XIST allele is transcriptionally inactivated.

Since the site of the mutation as well as several regulatory elements identified by in vitro studies may be involved in regulating Xist expression, examination of factor binding to these elements in vivo on both the active and inactive X chromosomes may provide insight into the molecular mechanisms controlling XIST gene transcription. Therefore, we used ligation-mediated PCR (LMPCR) in vivo footprinting with dimethyl sulfate (DMS), to examine the differential binding of transcription factors to the normal and mutant XIST alleles on the active and inactive X chromosomes.

Materials and Methods

Cell Lines

Four somatic cell hybrid lines, 4.12, 8121, a60-1A, and t60-17A2 were used for in vivo footprint analysis. Cell line 4.12 is a hamster-human hybrid containing only the active human X chromosome in the HPRT-deficient hamster cell line RJK88. Cell line 8121 is a hamster-human hybrid containing only the inactive human X chromosome in RJK88. 4.12 cells were cultured in Dulbecco's modified Eagle's medium (D-MEM) supplemented with hypoxanthine, aminopterin and thymidine (HAT), and 8121 cells were cultured in D-MEM.

To examine the mutant XIST gene, two human-mouse somatic cell hybrid lines, the a60-1A and the t60-17A2 cell lines, were prepared using pushmi-pullyu hybrid selection (17). The t60-17A2 cell line harbors the mutant XIST gene on an inactive human X chromosome which was maintained by growth at the non-permissive temperature of 39°C. The a60-1A cell lines harbors the mutant XIST gene on an active human X chromosome which was maintained by growth at 37°C in HAT supplemented media. Both cell lines were grown in a-MEM media supplemented with 10% FBS, L-glutamine and P/S.

Dimethyl Sulfate Treatment and DNA Isolation

Cells were treated with 0.5% DMS for 60, 90, 120, and 180 seconds, and DNA was isolated from treated cells as previously described (87). In vitro DMS treatment of purified DNA was performed as previously described (87).

Ligation-Mediated PCR (LMPCR) and Detection of In Vivo Footprints LMPCR was carried out as described by Litt et al. (87) For LMPCR, the lower strand was analyzed with primer set X (X1b:GATTTACTCAATGTCCCCATG, X2A:CTGGAAAACCCATTGAAGTTGTGACTCCTGG) and the upper strand was analyzed with primer set Y (Y1b:GGTGTTGGGGGAGTAGAAAA, Y2A:ACCCCAAGTGCAGAGAGATCTTCAGTCAGG). Reaction conditions for first strand synthesis and ligation were identical to those previously described (87). PCR conditions were the same as previously described (87) except samples were cycled in an MJ PTC-100 programmable thermal controller as follows: denatured for 95°C for 3 minutes then cycled 20 times at 95°C for 20 seconds, 64°C for 1 minute, and 76°C for 1 minute plus 5 seconds each cycle. This was followed by an additional 5 cycles at 95°C for 20 seconds, 64°C for 1 minute and 76°C for 2 minutes 40 seconds and a 10 minute extension at 76°C.

Gel electrophoresis and electroblotting were carried as described (87). The final DNA sequencing ladders were visualized using single-stranded hybridization probes synthesized from M13 clones containing the human XIST promoter region. Probe synthesis, hybridization, washing and autoradiography were performed as previously described (60, 63).

Results

To examine the differential binding of sequence-specific regulatory factors to the 5' region of the active and inactive minimal human XIST promoter, positions -3 to -173 of the upper strand and positions +33 to -115 of the lower strand (Figure 6.1) are examined by LMPCR-assisted DMS in vivo footprinting (DMS reacts preferentially with guanine nucleotides). Both the normal and mutant human XIST promoters are examined. This region includes the minimal promoter defined by deletion analysis, and the 5 critical regions identified by saturation mutagenesis and point mutation studies (56).

In Vivo Footprinting of the Normal Minimal XIST Promoter

Two human-hamster hybrid cell lines, 4.12 and 8121, are used for in vivo footprint analysis of the normal promoter on the active and inactive X chromosomes, respectively. Representative autoradiograms from analysis of the human XIST 5' region in these cells are shown in Figures 6.2 and 6.3. In Figure 6.2, the upper strand of the human XIST gene from positions -33 to -173 is analyzed with primer set Y. This region contains a potential TATA-binding protein (TBP) binding site, two potential Sp1 binding sites, and the string of 7 consecutive cytosines. Comparison of the DMS modification and cleavage pattern between naked DNA (Fig. 6.2, lanes 1 and 3) and in vivo treated DNA (Fig. 6.2, lanes 2 and 4) reveals footprinted nucleotides at positions -75 and -70 in the 8121 cell line which contains the active human XIST allele. These footprints occur within a potential Sp1 binding site which was shown to be important for XIST expression by in vitro saturation mutagenesis studies (56). No other footprints are observed on the upper strand of the normal active XIST allele in this region. In addition, analysis with

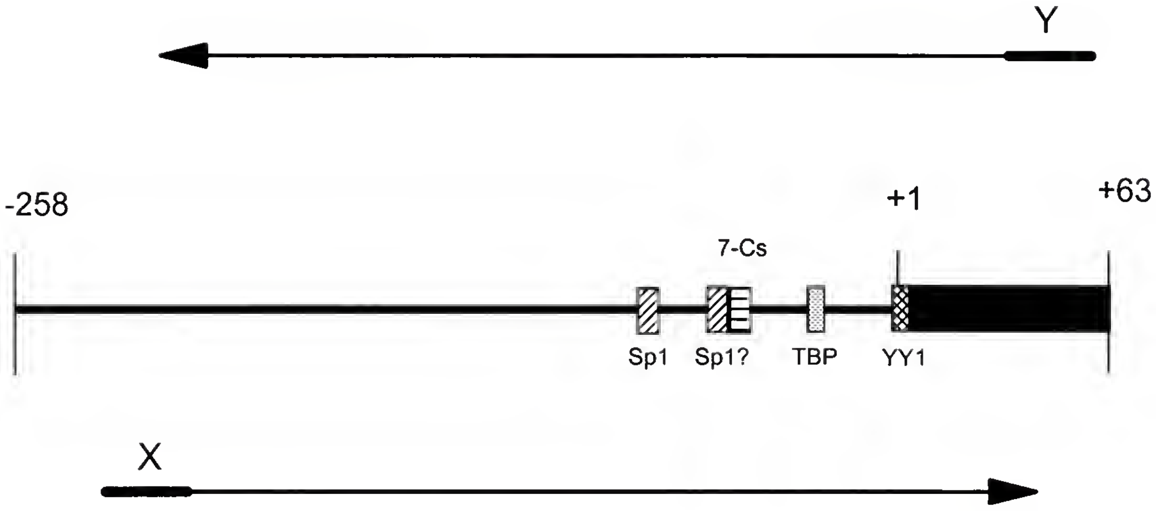
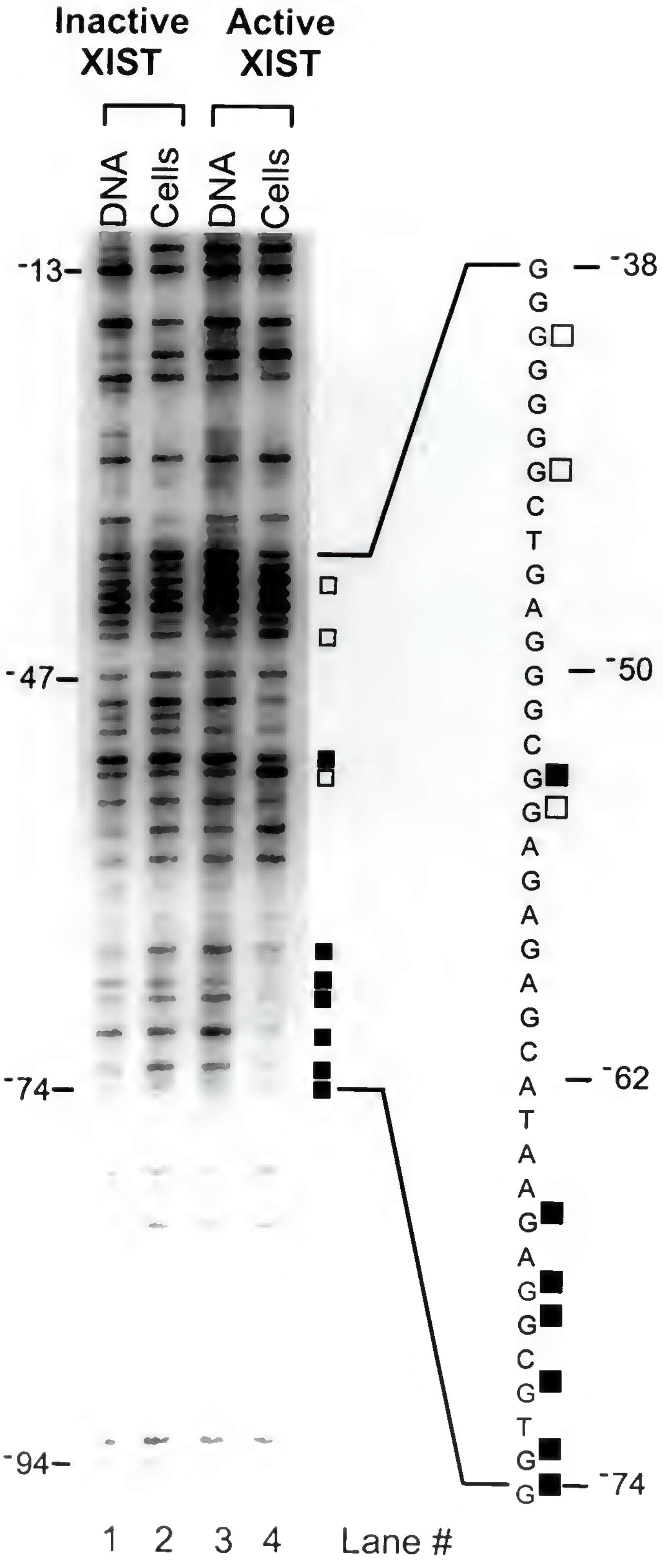


FIGURE 6.1 - Location of LMPCR primers for analysis of the human XIST 5' region. The center line represents the human XIST 5' regions with positions numbered relative to the transcription initiation site. The large filled rectangle that begins at position +1 indicates the transcribed portion. The smaller striped, stippled, and cross-hatched boxes indicate transcription factor binding sites. 7-Cs indicates a string of 7 cytosines which is footprinted and contains the C to G mutation. The horizontal arrows above and below the line indicate the positions of the primers sets used for LMPCR and the direction and approximate region examined. Primer set Y was used to examine the upper strand and primer set X was used to examine the lower strand.

FIGURE 6.2 - In vivo footprint analysis of the lower strand of the normal XIST promoter using primer set X. Autoradiogram of the guanine-specific sequencing ladder from DMS treatment. The nucleotide sequence of the region is shown to the right of the guanine-specific sequencing ladder generated by DMS; the position of each nucleotide is indicated relative to the transcription initiation site. Numbers to the left of the autoradiogram indicate positions of bands relative to the transcription initiation site. Open squares to the right of the nucleotide sequence and the autoradiogram represent sites of enhanced DMS reactivity, and solid squares represent guanine nucleotides protected from DMS modification. Lane numbers are designated at the bottom of the autoradiogram. DNA, denotes naked DNA purified from either 8121 or 4.12 cells and treated in vitro with DMS. Cells, denotes DNA from intact 8121 or 4.12 cells treated in vivo with DMS. Inactive XIST, denotes samples from the 4.12 cells which contain an inactive XIST allele on an active X chromosome. Active XIST, denotes samples from the 8121 cells which contain an active XIST allele on the inactive X chromosome.


Inactive XIST
Active XIST
DNA
Cells
DNA
Cells
-13
-47
-74
-94
1 2 3 4 Lane #
-38
-50
-62
-74
G G G G G G G C T G A G G G C G G A G A G A G C A T A A G A G G C G T G G

primer set Y reveals no footprints in the 4.12 cell line which contains the inactive human XIST allele.

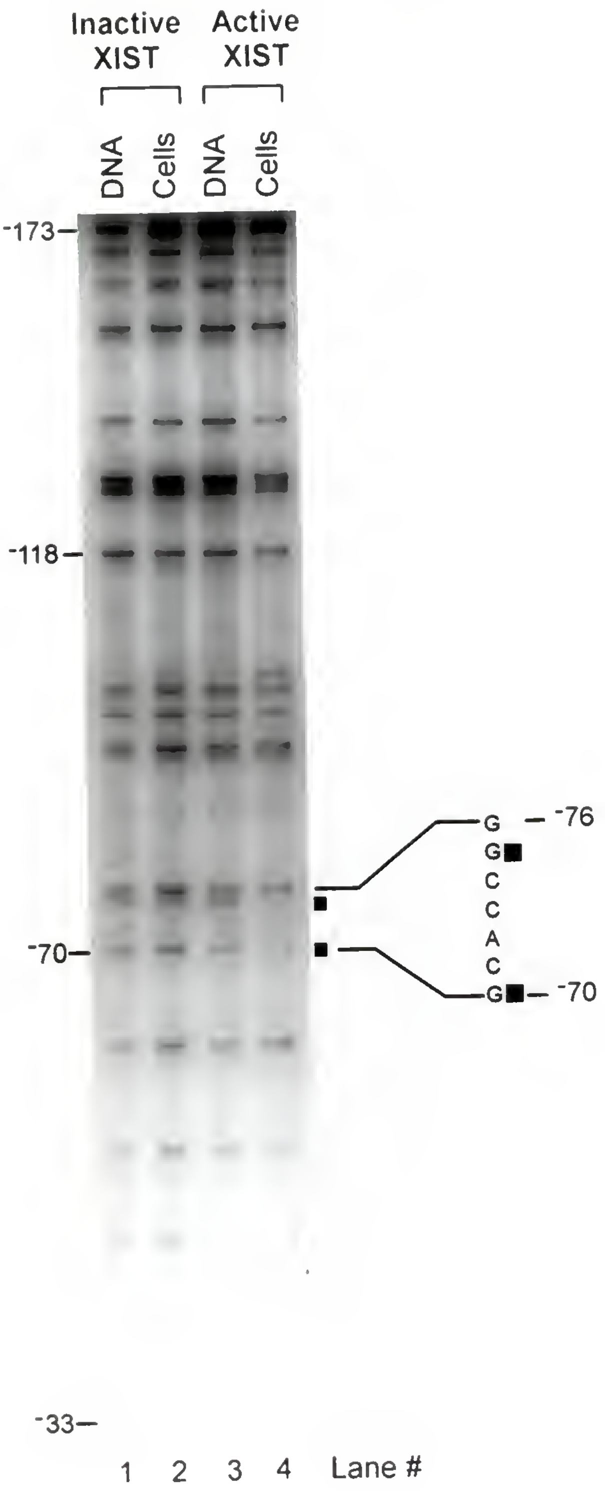
In Figure 6.3, the lower strand from positions -1 to -115 is analyzed with primer set X. Comparison of the DMS modification and cleavage pattern between naked DNA (Fig. 6.3, lanes 1 and 3) and in vivo treated DNA (Fig. 6.3, lanes 2 and 4) reveals several footprinted regions in samples containing the active XIST allele (Fig. 6.3, lane 4). Footprints are observed at positions -40, -44, -53, -54, -56, -66, -69, -71, -73, and -74 which are located within the two potential Sp1 binding sites (-49 to -56, and -66 to -75) and within the 7 consecutive cytosines (-38 to -44). The footprinted regions consist of DMS-protected sites as well as sites of enhanced DMS reactivity. No footprints are observed within this region in samples containing an inactive XIST allele (Fig. 6.3, lane 2).

For the region examined, all observed footprints occur on the active allele and no footprints are detected on the inactive allele. This pattern of differential transcription factor binding on the active and inactive alleles is similar to that observed in in vivo footprinting studies of the human and mouse HPRT gene (63, 87) and the X-linked human PGK-1 gene (119, 122). No in vivo footprints are detected over the TBP and YY1 binding sites in the active XIST promoter, most likely because these interactions do not involve guanine contacts.

In Vivo Footprinting of the Mutant XIST Promoter

Recently, Naumova et al. (109) reported a family which showed skewed X inactivation among its female members. Sequence analysis of the XIST gene promoter in this family revealed a C to G transversion at position -43, with the X chromosome

FIGURE 6.3 - In vivo footprint analysis of the upper strand of the normal XIST promoter using primer set Y. All symbols and designations are identical to those described in Fig. 6.2.

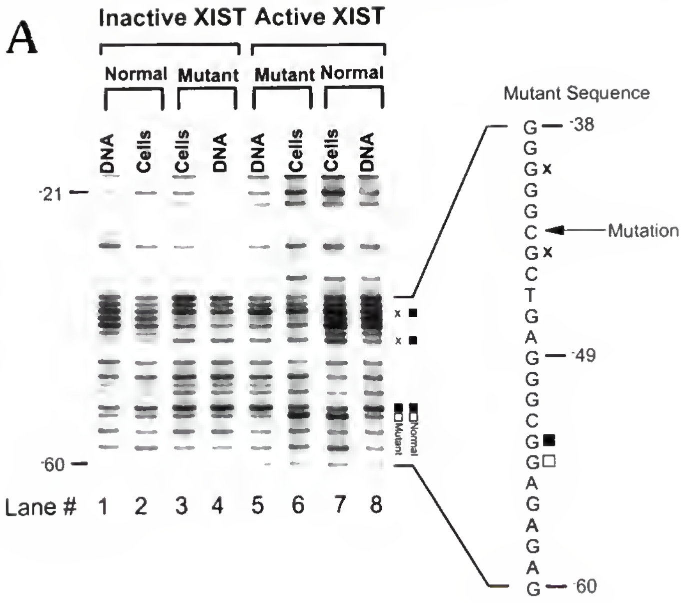

Inactive XIST
Active XIST
DNA
Cells
DNA
Cells
-173
-118
-70
-33
G
G■
C
C
A
C
G■
-76
-70
■
■
1
2
3
4
Lane #

carrying the mutant XIST promoter being preferentially inactivated. Furthermore, this mutation reduced XIST promoter activity 5-fold in transient expression assays (56). To assess the effect of this mutation on transcription factor binding in vivo, we performed LMPCR-assisted DMS in vivo footprinting of the mutant promoter on the active and inactive X chromosomes.

Two human-mouse hybrid cell lines, a60-1A and t60-17A2, each containing a single human X chromosome harboring the mutant XIST promoter in a mouse cell background, are used in the footprinting studies. The a60-1A cell line contains an active mutant X chromosome, and the t60-17A2 cell line contains an inactive mutant X chromosome.

Representative autoradiographs from analysis of the mutant XIST promoter in these cell lines are shown in Figures 6.4 and 6.5. Figure 6.4a and b show the analysis of the lower strand from positions -21 to -94. Comparison of the normal and mutant alleles reveals identical footprints over the two potential Sp1 sites at positions -49 to -56 and -66 to -75 (lanes 4-8). However, at positions -40 and -44 where the normal active XIST allele exhibits footprints of enhanced DMS reactivity (Fig. 6.4a, lane 7; and Fig. 6.3, lane 4), the mutant allele lacks these footprints (Fig. 6.4a, lane 6). This loss of footprints in the mutant allele occurs in the immediate vicinity of the mutation (position -43). Thus, in the mutant allele, the two in vivo footprints occurring over sites outside of the mutation remain unchanged relative to the normal allele, while footprints near the mutation disappear. No new footprints appear on the lower strand of the mutated active allele (Fig. 6.4a and b, lane 6), and no footprints are observed on the mutated inactive promoter (Fig. 6.4a and b, lane 3).

FIGURE 6.4 - In vivo footprint analysis of the lower strand of the mutant XIST promoter using primer set X. Autoradiograms of the guanine-specific sequencing ladders from position -21 to -60 (A) and from positions -66 to -94 (B). Symbols and designations are identical to those described in Fig. 2. Small x's to the right of the sequence denote guanines which are not footprinted. The arrow designated Mutation to the right denotes the C to G transversion at base -43 within the XIST promoter. Normal denotes samples from either the 4.12 or 8121 which contain a normal XIST promoter. Mutant denotes samples form either the a60-1A or t60-17A2 cells which contain a mutant XIST promoter. DNA, denotes naked DNA purified from either 4.12, 8121, a60-1A or t60-17A2 cells and treated in vitro with DMS. Cells, denotes naked DNA from intact 4.12, 8121, a60-1A or t60-17A2 cells treated in vivo with DMS. Inactive XIST, denotes samples from the a60-1A or 4.12 cells which contain an inactive XIST allele. Active XIST, denotes samples from the t60-17A2 or 8121 cells which contain an active XIST allele.


A
Inactive XIST
Active XIST
Normal
Mutant
Mutant
Normal
DNA
Cells
Cells
DNA
DNA
Cells
Cells
DNA
-21
-60
Lane # 1 2 3 4 5 6 7 8
x ■
x ■
■ Mutant
□ Normal
Mutant Sequence
G — -38
G
G x
G
G
C ← Mutation
G x
C
T
G
A
G — -49
G
G
C
G ■
G □
A
G
A
G
A
G — -60

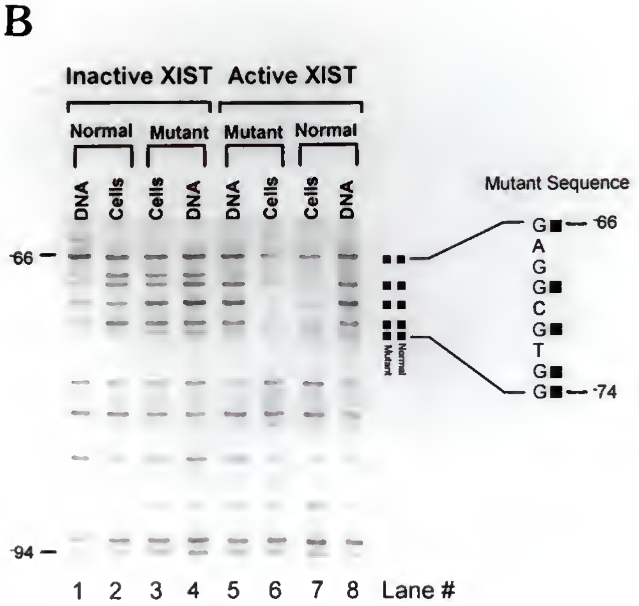

B
Inactive XIST
Active XIST
Normal
Mutant
Mutant
Normal
DNA
Cells
Cells
DNA
DNA
Cells
Cells
DNA
-66
-94
1 2 3 4 5 6 7 8 Lane #
■ Mutant
■ Normal
Mutant Sequence
G ■ — -66
A
G
G ■
C
G ■
T
G ■
G ■ — -74

FIGURE 6.5 - In vivo footprint analysis of the upper strand of the mutant_XIST promoter using primer set Y. All symbols and designations are identical to those described in Fig. 6.4.

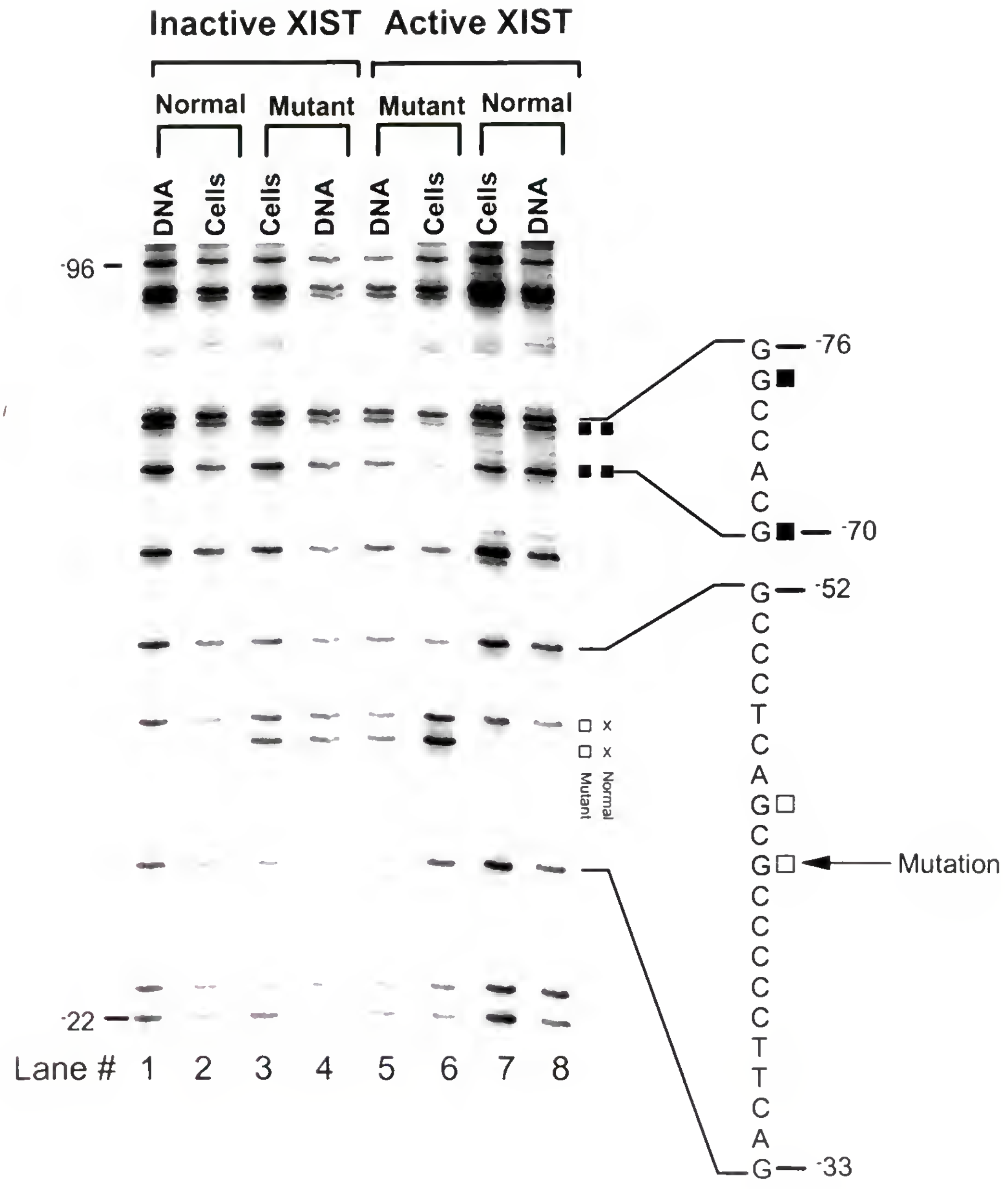

Inactive XIST
Active XIST
Normal
Mutant
Mutant
Normal
DNA
Cells
Cells
DNA
DNA
Cells
Cells
DNA
-96
-22
Lane # 1 2 3 4 5 6 7 8
x Normal
□ Mutant
G -76
G
C
C
A
C
G -70
G -52
C
C
C
T
C
A
G
C
G
C
C
C
C
C
T
T
C
A
G -33
Mutation

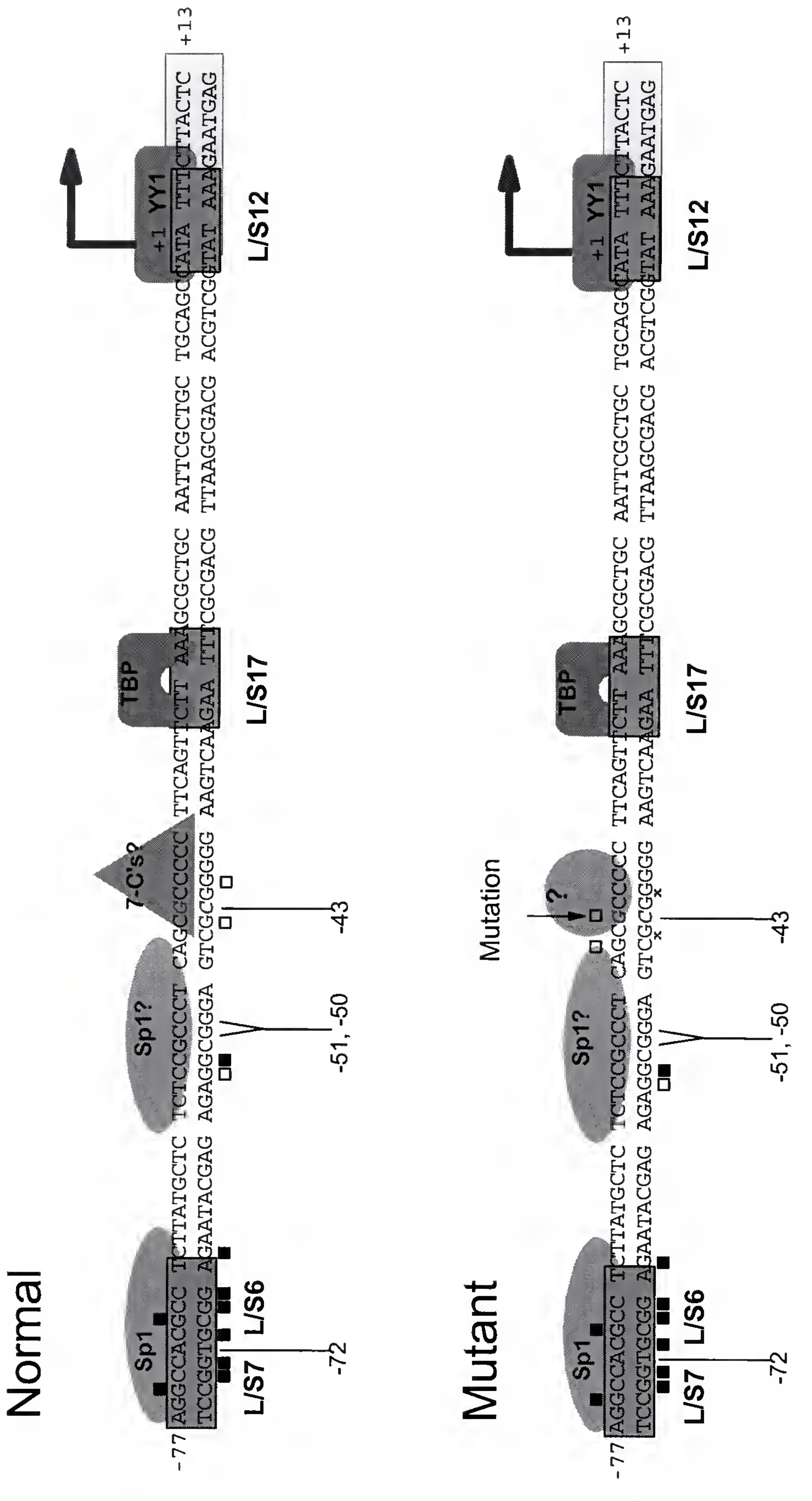
Figure 6.5 shows analysis of the upper strand from positions -22 to -96 of both the normal and mutant XIST promoter. The two footprinted nucleotides in the normal active allele at positions -75 and -70 (Fig. 6.5, lane 8; and Fig. 6.2, lane 4) are also present in the mutant active allele (Fig. 6.5, lane 6). However, examination of the active mutant promoter reveals two new footprinted sites at positions -43 and -45 (Fig. 6.5, lane 6). These new footprints appear as bands of increased intensity (relative to the naked DNA control in Fig. 6.5, lane 5), reflecting sites of enhanced DMS reactivity. The footprint in the active mutant allele at position -43 occurs at the mutated base within the normal string of seven cytosines; the C to G transversion appears as a new guanine band at position -43 in the sequencing ladder of samples containing a mutant XIST promoter (Fig. 6.5, lanes 3 to 6). The new footprint at position -45 occurs immediately 5' of the seven consecutive cytosines. Both of these footprints are absent in the normal active and inactive alleles (Fig. 6.5, lanes 2 and 7), as well as the mutant inactive allele (Fig. 6.5, lane 3). Thus, the mutant active XIST promoter exhibits an altered (but not deleted) in vivo footprint pattern on the upper strand relative to the normal active promoter at and adjacent to the site of the C to G transversion at position -43.

Discussion

In Vivo Footprinting

Figure 6.6 summarizes results of in vivo footprinting studies of the normal and mutant XIST promoters, and includes results of in vitro characterization of the human XIST promoter by Hendrich et al. (56). For the normal human XIST gene at least three regions show in vivo footprints. Of these three sites, in vitro studies and sequence analysis suggest that two of the sites are potential Sp1 binding sites. The third site

Figure 6.6 - Summary of in vivo footprints in the normal and mutant XIST promoter on the inactive X chromosome. The sequence is numbered relative to the transcription start site which is denoted by a +1. The large arrow indicates the direction of transcription. The striped rectangle indicates transcribed sequence. Normal designates the normal XIST promoter. Mutant designates the XIST promoter harboring the C to G mutation at -43 which is indicated by bold italic letters. Footprinted nucleotides are denoted with small boxes. Small filled boxes denote guanine residues with decreased DMS modification (protections). Small open boxes denote guanine residues with increase DMS modification (enhancements). The boxed and shaded sequence indicates the position of linker scanner mutations which decreased XIST expression (51). Numbered vertical lines indicate base mutated in vitro by Hendrich et al. (51) that decreased XIST expression. The small x's denote guanines which are not footprinted. Sp1, TBP, and YY1 indicate potential transcription factors and are designated with large shaded shapes.


Normal
-77
AGGCCACGCC TCTTATGCTC TCTCCGCCCT CAGCGCCCCC TTCAGTTCTT AAAGCGCTGC AATTCGCTGC TGCAGCCATA TTTCTTACTC
TCCGGTGCGG AGAATACGAG AGAGGCGGGA GTCGCGGGGG AAGTCAAGAA TTTCGCGACG TTAAGCGACG ACGTCGGTAT AAAGAATGAG
+13
Sp1
Sp1?
7-C's?
TBP
+1
YY1
L/S7
L/S6
L/S17
L/S12
-72
-51, -50
-43
Mutant
-77
AGGCCACGCC TCTTATGCTC TCTCCGCCCT CAGCGCCCCC TTCAGTTCTT AAAGCGCTGC AATTCGCTGC TGCAGCCATA TTTCTTACTC
TCCGGTGCGG AGAATACGAG AGAGGCGGGA GTCGCGGGGG AAGTCAAGAA TTTCGCGACG TTAAGCGACG ACGTCGGTAT AAAGAATGAG
+13
Sp1
Sp1?
Mutation
?
TBP
+1
YY1
L/S7
L/S6
L/S17
L/S12
-72
-51, -50
-43

consists of a string of seven cytosines (in the normal allele). All of the observed in vivo footprints are associated only with the transcriptionally active XIST allele. These results are consistent with transient expression and in vitro studies which indicated that five regions in the active XIST promoter were important for high levels of XIST expression (56), including the three regions containing in vivo footprints. The two remaining sites, the TBP and YY1 binding sites, did not display detectable in vivo footprints, probably because factor binding at these sites does not involve guanine contacts detected by DMS.

Comparison of the in vivo footprint patterns between the normal and mutant active alleles shows a clear difference in the footprint pattern. The footprints for the two Sp1 sites are identical in both the normal and mutant alleles. However, the footprint pattern over the string of seven cytosines is altered by the mutation (which occurs within this sequence). On the lower strand, the two footprinted guanines which showed enhanced DMS reactivity in the normal promoter (at positions - 40 and - 44) are no longer footprinted in the mutant (denoted by x's on the mutant sequence in Figure 6.6). Furthermore, two new footprints appear on the upper strand of the mutant at positions -45 and -43 (denoted by open boxes on the mutant sequence in Figure 6.6). These changes in the footprint pattern of the mutant occur in the immediate vicinity of the C to G mutation at position -43. In fact, the new guanine nucleotide in the mutant is itself footprinted.

The notable alteration of the in vivo footprint pattern in the immediate region of the mutation suggests that a different DNA-binding factor is interacting with the mutated sequence. The change in the footprints over this site and the skewed pattern of X inactivation associated with this mutation suggest that this site may be involved in selecting which XIST allele is transcriptionally active or inactive. Alternatively, it is possible that the same DNA-binding factor is interacting with this site in both the normal and mutant alleles, but has merely changed its contact nucleotides because of the mutation. This latter possibility seems unlikely given the extent of the changes in the guanine contact pattern (on both the upper and lower strand) that constitute the normal and mutant in vivo footprints.

Implications for X Inactivation

The selective inactivation of the X chromosome harboring the mutation in this family suggests that this site may be involved in determining which XIST allele is active or inactive prior to the onset of X inactivation. The fact that this mutation reduces XIST expression 5-fold in transient expression assays suggests that the level of XIST expression alone cannot account for selection of the X chromosome that is to be inactivated. However, gene knock-out studies in mice indicate that some XIST expression in cis is necessary for X inactivation (118). Our in vivo footprinting results suggest a DNA-binding regulatory factor(s) which is involved in establishing the active and inactive XIST alleles may interact within the XIST minimal promoter, and that the mutation at position -43 alters the binding of regulatory factors to the XIST promoter. We propose that the mutation alters the normal process of marking the active and inactive XIST alleles by acting as the binding site for a different regulatory factor, resulting in preferential silencing of the normal XIST gene. As a result, the X chromosome harboring the mutation is more likely to express XIST and be inactivated.

CHAPTER 7
CONCLUSIONS AND FUTURE DIRECTIONS

In this dissertation, I have examined several mechanisms that are involved in the differential regulation of individual X-linked genes by X chromosome inactivation. In chapter 2, the differential binding of regulatory factors and the high-resolution methylation pattern of the mouse HPRT gene on the active and inactive X chromosome were investigated. In vivo footprinting analysis revealed several footprints on the active allele but no footprints on the inactive allele. A similar pattern of differential factor binding was observed on the human HPRT gene (63) and human PGK-1 gene (122). Careful comparison of the in vivo footprinting data among these genes revealed no unique transcription factor binding site common to all three genes that may be responsible for the coordinate differential regulation of X-linked genes. The absence of DNA-protein interactions on the inactive allele of the human HPRT and PGK-1 and the mouse HPRT gene suggests that a sequence-specific repressor is not responsible for the coordinate silencing of X-linked genes on the inactive X chromosome. Therefore, these studies argue against the hypothesis that a DNA sequence that binds either an activator or repressor protein within the promoter of X-linked genes subject to inactivation is responsible for the process of X inactivation. However, it is possible that such a factor was not detected by these studies or is located outside of the regions examined.

The absence of any factors bound to the inactive allele of the the mouse HPRT (Chapter 2), and the human HPRT (63) and PGK-1 (119, 122) genes suggested that the inaccessibility of the inactive allele may be controlled by another mechanism such as DNA methylation and/or chromatin structure. For this reason the high-resolution DNA methylation pattern was also examined (Chapter 2). Genomic sequencing of the mouse HPRT gene indicated that all of the examined CpG dinucleotides on the active allele are unmethylated, whereas the majority of the CpG's are methylated on the inactive allele, with a few partial and unmethylated sites scattered throughout the promoter. Comparison of the methylation data and footprinting data revealed no discernable correlation among methylated sites and unmethylated sites with in vivo footprints. This differs from the methylation pattern of the human HPRT gene, which shows a distinct region of hypomethylation over the GC boxes (63). Considering this, and the fact that Sp1 binding is not affected by DNA methylation (54, 57), it is probable that methylation of the inactive promoter is not directly interfering with transcription factor binding. This hypothesis is supported by the methylation pattern of the mouse PGK-1 gene, which has only one fully methylated site on the inactive allele (143). Other evidence supports a role for the methyl binding proteins (MeCp1 and MeCp2) interacting the methylated promoter of inactive X-linked genes (8, 9). However, we did not detect any factors bound to the inactive allele over methylated or unmethylated sites. This leads us to conclude that MeCp1 or MeCp2 are not be involved in the maintenance of X-linked genes subject to transcriptional silencing on the inactive X chromosome. Besides the hypermethylation of the inactive promoter relative to the active promoter, comparison of the methylation patterns of the human and mouse HPRT and PGK-1 genes failed to reveal an obvious methylation pattern that could be correlated with transcriptional silencing of the inactive allele. These experiments indicate that although DNA methylation is probably involved in X inactivation, the exact function is still unclear. This leads us to believe that DNA methylation may function collaboratively with other regulatory mechanisms such as chromatin structure to control the accessibility of transcription factors to X-linked promoters on the inactive X chromosome.

In vivo footprinting experiments of the human and mouse HPRT genes revealed a novel 9-bp sequence that was identical in both genes and footprinted at the same guanine residue (63, 87). Chapter 3 discusses the reconstitution experiments of the -91 footprinted, which occurs within the novel nine base pair sequence using gel mobility-shift assays. The results revealed the formation of multiple DNA-protein complexes. Of these complexes, specific competitors suggested that DNA-protein complex I and III may be represent reconstitution of the DNA-protein interactions occurring at the -91 guanine in vivo. However, these experiments need to be repeated to strengthen this hypothesis. The facts that the novel 9-bp sequence shows a DNA-protein complex in vitro and that this sequence is footprinted in vivo in both the human and mouse active HPRT alleles argue strongly that potentially some unique factor(s) is binding to this site. Since this novel nine base sequence is located immediately 3 prime of the transcription initiation sites in the human and mouse HPRT gene, it is possible that the factor interacting with this site represent a new transcriptional initiator.

Chapter 4 discusses the time course of events following the reactivation of the human HPRT gene with the demethylating agent 5aCdr. These experiments coupled with previous studies by Sasaki et al. (133) suggested that hemi-demethylation and changes in chromatin structure precede the appearance of symmetrical demethylation, transcription factor binding, and synthesis HPRT mRNA. These experiments suggest that changes in chromatin structure occur independent of transcription factor binding. From these data it appears that reactivation of the HPRT genes first requires hemi-demethylation and opening of chromatin structure before transcription factors are capable of binding the promoter and initiating transcription of the HPRT gene. The results of this experiment support the function of DNA methylation and demethylation primarily affecting chromatin structure. Subsequently, as a result of changes in chromatin structure, these changes influence sequence-specific DNA-protein interactions in the promoter region. Therefore, we hypothesize that chromatin structure is playing an important role in regulating the differential expression of X-linked genes caused by X inactivation.

Chapter 5 discusses the timing of inactivation of the X chromosome during spermatogenesis as determined by examining the binding of transcription factors to the mouse HPRT promoter at different stages of spermatogenesis. These studies suggest that the bulk of transcriptional silencing of the male X chromosome occur just prior to or during the late pachytene spermatocyte stage of spermatogenesis.

Chapter 6 discusses the differential binding of transcription factors to the active and inactive human XIST allele. In these studies, both a mutant and normal human XIST alleles were examined. In vivo footprinting analysis of the normal XIST allele revealed factors bound to the active XIST allele and no factors bound to the inactive XIST allele, similar to my observations of the mouse HPRT gene. Only one of the footprinted region occurred over a novel sequence, a string of seven cytosines. The remaining footprinted regions occurred over potential Sp1 sites. These data suggest that mechanisms such as DNA methylation and/or chromatin structure may be involved in controlling the differential regulation of the human XIST in somatic cells. In vivo footprinting of the mutant XIST gene showed a different footprint pattern than the normal allele. In fact, new footprints were observed over the mutated site. Since patients harboring X chromosomes with the mutant XIST promoter show skewed inactivation, in which the mutant X chromosome is selectively inactivated, it is possible that the propensity to establish expression of the mutant XIST allele, and thereby inactivate the corresponding X chromosome, may be caused by the disruption of transcription factor binding at this site. Perhaps a new/different transcription factor is binding the mutant and effect XIST expression.

In summary, my dissertation project involved investigating several mechanisms of regulation of the transcription of X-linked genes on the active and inactive X chromosomes. These studies suggest that chromatin structure may be playing an important role in the process of X inactivation. Therefore, studies which examine the mechanisms by which alterations of chromatin structure control gene expression need to be examined for X-linked genes subject to inactivation. A better understanding of chromatin structure will certainly reveal some of the mechanisms involved in the process of X inactivation.

Implications for X Inactivation

It has been postulated that each X-linked gene subject to inactivation is regulated by the differential binding of repressor and activator proteins that interact with a cis-acting regulatory element(s) near or within the promoter of each X-linked gene subject to inactivation (94). In vivo footprinting studies of the human PGK-1 and HPRT genes and the mouse HPRT gene revealed several DNA-protein interactions on the active allele (Figure 7.1). The majority of the transcription factors observed were ubiquitous factors such as Sp1, NF-κB, or AP-2. A DNA regulatory protein involved in the coordinate transcriptional activation or silencing of individual X-linked genes on the active and inactive X chromosome will probably be unique and common to all X-linked genes subject to regulation by X chromosome inactivation. Therefore, ubiquitous factors are unlikely to be the primary elements involved in the coordinated regulation of the X-linked genes. Novel binding sites were observed in all three genes. However, there is no unique DNA-protein interaction common to the three X-linked genes that have been in vivo footprinted. Consequently, it is unlikely that a transcriptional activator, which binds to the promoter of individual X-linked genes, and is responsible for their coordinate activation, is involved. Furthermore, the absence of in vivo footprints on the inactive allele argues against a transcriptional repressor binding to the promoter and being responsible for the coordinate inactivation of individual X-linked genes. This suggests that another mechanism, such as DNA methylation and/or chromatin structure, may be involved in the coordinate regulation of X-linked genes that are subject to transcriptional regulation by X inactivation.

Several experiments have demonstrated that there is a correlation between hypermethylation of the inactive X-linked promoters and transcriptional silence. This is particularly true for housekeeping genes. Genomic sequencing studies have examined the

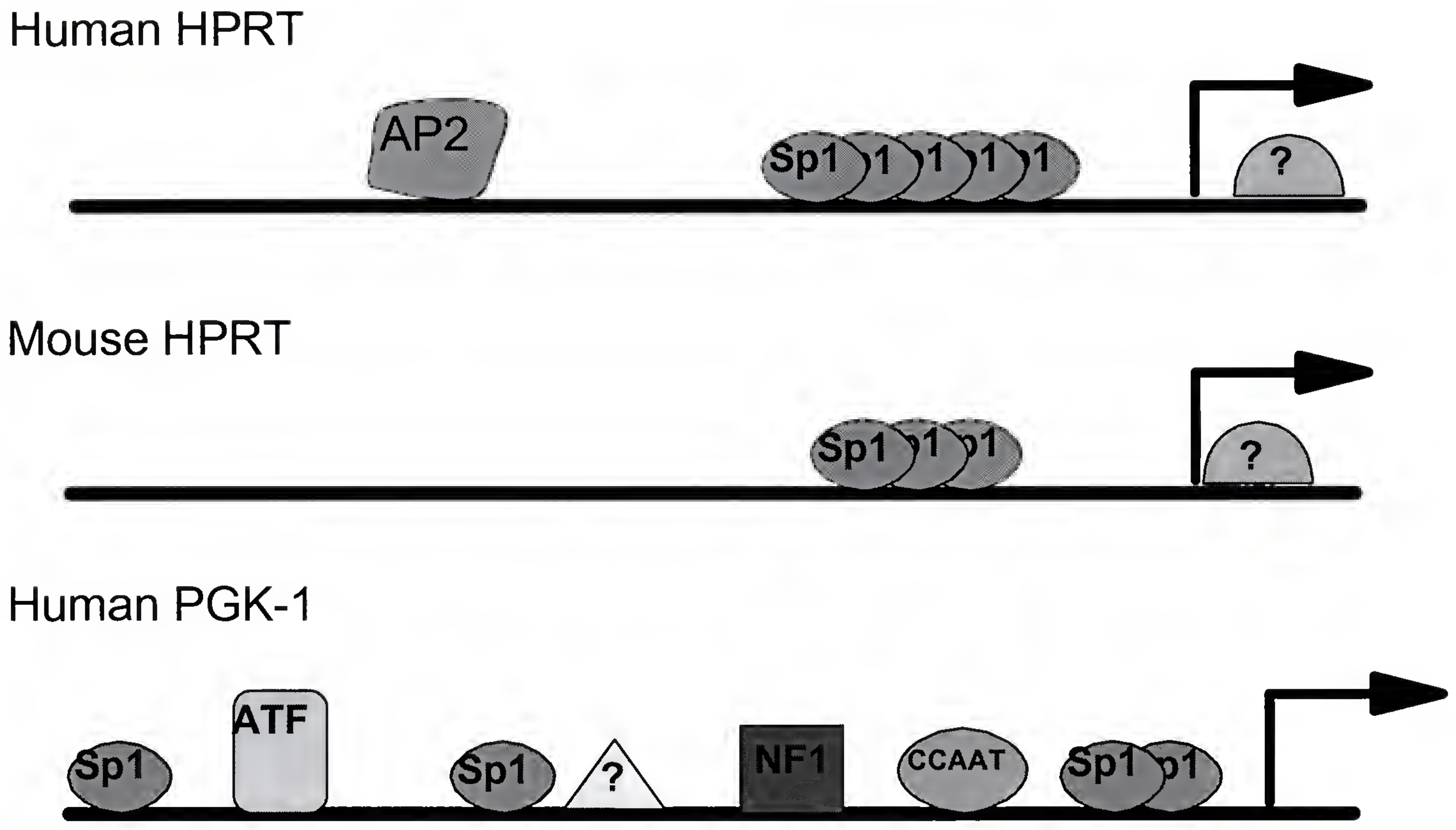


Figure 7.1 - Comparison of in vivo footprints from the active allele of 3 X-linked genes. The thick line indicates either the human HPRT, mouse HPRT or the human PGK-1 gene. The large arrow designates the approximate site of transcription initiation. Sshaded shapes designate the approximate position of potential transcription factors that were observed by in vivo footprinting experiments. Unlabeled shapes(?) are novel transcription factors.

Gene	Active Allele	Inactive Allele
Human HPRT	Unmethylated	Methylated at all CpG's except within GC boxes
Mouse HPRT	Unmethylated	Methylated at most CpG's; unmethylated CpG's random
Human PGK-1	Unmethylated	Methylated at all CpG's
Mouse PGK-1	Unmethylated	Only 1 fully methylated CpG; others unmethylated or partially methylated

Figure 7.2 - The methylation patterns of the promoter of four X-linked housekeeping genes

high resolution methylation pattern of the 5' region of the human and mouse PGK-1 and HPRT alleles on the active and inactive X chromosomes (Figure 7.2). For all four genes, every cytosine examined was unmethylated on the active allele. Comparison of the methylation pattern of the inactive allele did not reveal a common discernible pattern that may be responsible for the transcriptional silencing of the inactive allele. Furthermore, for the human PGK-1 and mouse HPRT genes there was no correlation between the patterns of DNA methylation and the in vivo footprinting pattern. This will probably also be true for the mouse PGK-1 gene, which has only one methylated site on the inactive allele. This suggests that DNA methylation is unlikely to silence X-linked genes transcriptionally by directly inhibiting transcription factor binding. Bird et al. (84) have proposed that a methyl binding protein, such as MeCp1 and MeCp2, binds to methylated DNA and prevent transcription factor binding. However, in vivo footprinting studies have not detected any such factors bound to the inactive allele. Therefore, if DNA methylation is involved in the differential accessibility of the active and inactive X chromosomes to transcription factors, it is probably acting in coordination with another mechanism such as chromatin structure.

This hypothesis is supported by the temporal events that follow reactivation of the human HPRT gene with the demethylating agent 5aCdr. In these experiments (Chapter 4), the following sequence of events occurred in order after treatment with 5aCdr: 1) the chromatin became more accessible to nucleases and the DNA became hemi-demethylated on the inactive HPRT promoter; and 2) transcription factors bound to the HPRT promoter, the DNA became fully demethylated, and the

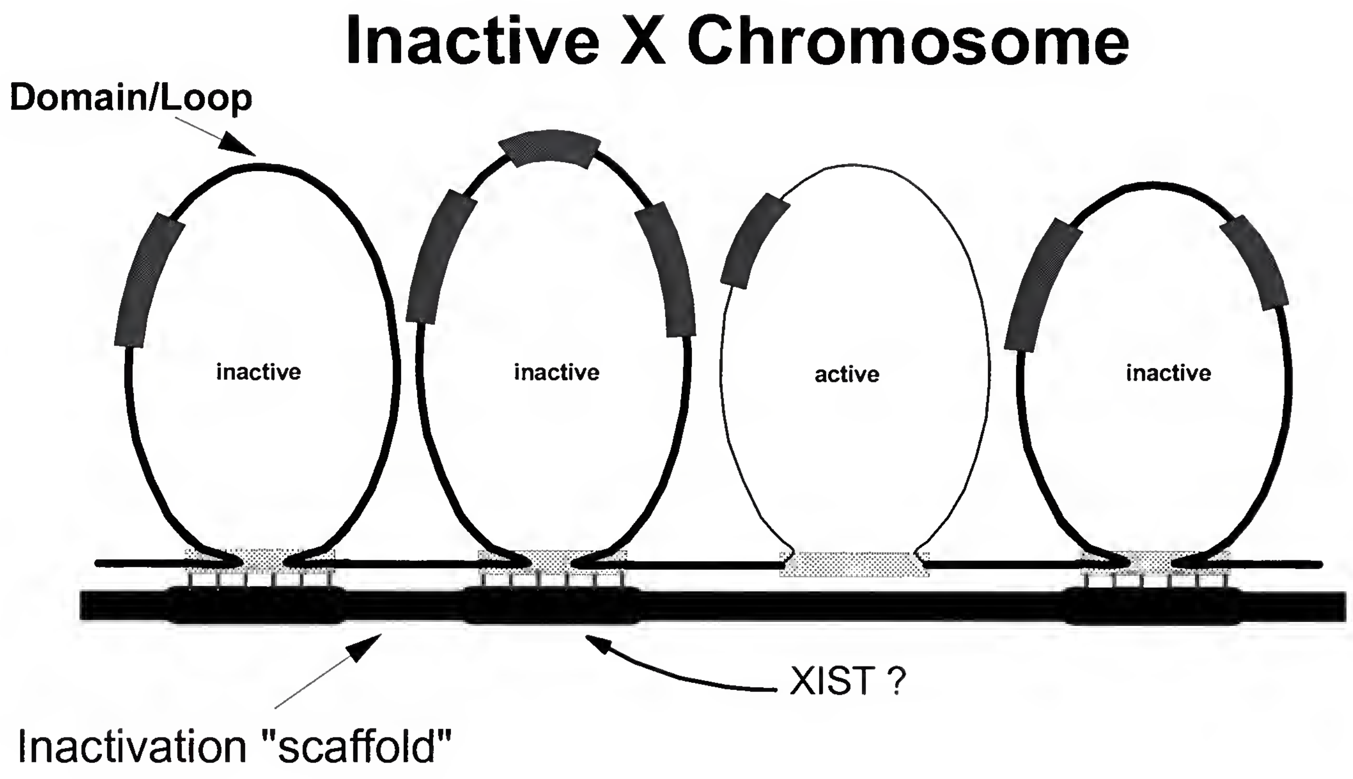


FIGURE 7.3 - Model for the transcriptional regulation of X-linked genes by chromatin domains. One or more X-linked genes (shown as boxes on the loops) are contained within discrete chromatin domains (shown as loops). Each domain contains regulatory elements that act to control the transcriptional activity of the genes contained within the domains. In this case an inactivation scaffold interacts with regulatory elements within each domain and thereby transcriptionally silences the genes within the domains that interact with the inactivation scaffold. Genes which escape inactivation might be located in active domains. An active domain would not contain the elements that interact with the inactivation scaffold to signal transcriptional silencing.

HPRT mRNA was transcribed. This temporal set of events suggests that DNA methylation is acting by locking in a chromatin architecture that is inaccessible to transcription factor binding. Furthermore, this suggests that the primary mechanism that is involved in the differential regulation of X-linked genes differences in chromatin structure.

Evidence suggests that one or more eukaryotic genes are organized into discrete functional units and/or chromatin domains (3, 47, 55, 65, 79, 149) and that these domains contain elements invovled in their independent regulation (21, 31, 68, 78, 83, 88, 98, 125, 139, 149). If chromatin structure is involved in the differential regulation of individual X-linked genes, it is possible that mechanisms by which X chromosome inactivation silences X-linked genes is occurring within these domains. One can imagine a system of coordinate regulation in which X-linked genes are organized into individual domains (Figure 7.3). Interacting with critical regulatory elements within this domain is a scaffold that is involved in changing each chromatin domain's accessibility to transcription factors. The composition of this scaffold may be composed of the XIST transcript which has been shown to be involved in X chromosome inactivation. Once a domain is signaled to be inaccessible to transcription factors it may become methylated to lock it into that inaccessible conformation. Therefore, genes which are transcriptionally silenced on the inactive X chromosome would be contained in compact methylated chromatin domains which are inaccessible to transcription factors.

REFERENCE LIST

1. **Adler, D. A., S. L. Bressler, V. M. Chapman, D. C. Page, and C. M. Disteche.** 1991. Inactivation of the Zfx gene on the mouse X chromosome [published erratum appears in Proc Natl Acad Sci U S A 1991 Jul 1; 88(13):5937]. Proc. Natl. Acad. Sci. U.S.A. **88:**4592-4595.

2. **Agulnik, A. I., M. J. Mitchell, M. G. Mattei, G. Borsani, P. A. Avner, J. L. Lerner, and C. E. Bishop.** 1994. A novel X gene with a widely transcribed Y-linked homologue escapes X-inactivation in mouse and human. Hum. Mol. Genet. **3:**879-884.

3. **Alevy, M. C., M. J. Tsai, and O. M. BW.** 1984. DNase I sensitive domain of the gene coding for the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase. Biochemistry **23:**2309-14.

4. **Bellve, A. R., J. C. Cavicchia, C. F. Millette, D. A. O'Brien, Y. M. Bhatnagar, and M. Dym.** 1977. Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. J. Cell Biol. **74:**68-85.

5. **Ben Hattar, J., P. Beard, and J. Jiricny.** 1989. Cytosine methylation in CTF and Sp1 recognition sites of an HSV tk promoter: effects on transcription in vivo and on factor binding in vitro. Nucleic Acids Res. **17:**10179-10190.

6. **Bird, A., M. Taggart, M. Frommer, O. J. Miller, and D. Macleod.** 1985. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell **40:**91-99.

7. **Blake, M. C., R. C. Jambou, A. G. Swick, J. W. Kahn, and J. C. Azizkhan.** 1990. Transcriptional initiation is controlled by upstream GC-box interactions in a TATAA-less promoter. Mol. Cell. Biol. **10:**6632-6641.

8. **Boyes, J., and A. Bird.** 1991. DNA methylation inhibits transcription indirectly via a methyl- CpG binding protein. Cell **64:**1123-1134.

9. **Boyes, J., and A. Bird.** 1992. Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. EMBO. J. **11:**327-333.

10. **Brandeis, M., D. Frank, I. Keshet, Z. Siegfried, M. Mendelsohn, A. Nemes, V. Temper, A. Razin, and H. Cedar.** 1994. Sp1 elements protect a CpG island from de novo methylation. Nature **371:**435-438.

11. **Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian.** 1986. Purification and biochemical characterization of the promoter- specific transcription factor, Sp1. Science **234:**47-52.

12. **Brockdorff, N., A. Ashworth, G. F. Kay, P. Cooper, S. Smith, V. M. McCabe, D. P. Norris, G. D. Penny, D. Patel, and S. Rastan.** 1991. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. Nature **351:**329-331.

13. **Brockdorff, N., A. Ashworth, G. F. Kay, V. M. McCabe, D. P. Norris, P. J. Cooper, S. Swift, and S. Rastan.** 1992. The product of the mouse Xist gene is a 15 kb inactive X- specific transcript containing no conserved ORF and located in the nucleus. Cell **71:**515-526.

14. **Brown, C. J., A. Ballabio, J. L. Rupert, R. G. Lafreniere, M. Grompe, R. Tonlorenzi, and H. F. Willard.** 1991. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature **349:**38-44.

15. **Brown, C. J., B. D. Hendrich, J. L. Rupert, R. G. Lafreniere, Y. Xing, J. Lawrence, and H. F. Willard.** 1992. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell **71:**527-542.

16. **Brown, C. J., R. G. Lafreniere, V. E. Powers, G. Sebastio, A. Ballabio, A. L. Pettigrew, D. H. Ledbetter, E. Levy, I. W. Craig, and H. F. Willard.** 1991. Localization of the X inactivation centre on the human X chromosome in Xq13. Nature **349:**82-84.

17. **Brown, C. J., and H. F. Willard.** 1994. The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. Nature **368:**154-6.

18. **Brown, C. J., and H. F. Willard.** 1989. Noninactivation of a selectable human X-linked gene that complements a murine temperature-sensitive cell cycle defect. Am. J. Hum. Genet. **45:**592-598.

19. **Buschhausen, G., B. Wittig, M. Graessmann, and A. Graessmann.** 1987. Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene. Proc. Natl. Acad. Sci. U.S.A. **84:**1177-1181.

20. **Carthew, R. W., L. A. Chodosh, and P. A. Sharp.** 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell **43:**439-48.

21. **Caterina, J. J., D. J. Ciavatta, D. Donze, R. R. Behringer, and T. M. Townes.** 1994. Multiple elements in human beta-globin locus control region 5' HS 2 are involved in enhancer activity and position-independent, transgene expression. Nucleic Acids Res. **22:**1006-11.

22. **Cattanach, B. M.** 1975. Control of chromosome inactivation. Annu. Rev. Genet. **9:**1-18.

23. **Cattanach, B. M.** 1970. Controlling elements in the mouse X-chromosome. 3. Influence upon both parts of an X divided by rearrangement. Genet. Res. **16:**293-301.

24. **Cattanach, B. M.** 1974. Position effect variegation in the mouse. Genet. Res. **23:**291-306.

25. **Cattanach, B. M., and J. H. Isaacson.** 1967. Controlling elements in the mouse X chromosome. Genetics **57:**331-346.

26. **Cattanach, B. M., J. N. Perez, and C. E. Pollard.** 1970. Controlling elements in the mouse X-chromosome. II. Location in the linkage map. Genet. Res. **15:**183-195.

27. **Cattanach, B. M., C. E. Pollard, and J. N. Perez.** 1969. Controlling elements in the mouse X-chromosome. I. Interaction with the X-linked genes. Genet. Res. **14:**223-235.

28. **Cattanach, B. M., and C. E. Williams.** 1972. Evidence of non-random X chromosome activity in the mouse. Genet. Res. **19:**229-240.

29. **Chapman, V. M., P. G. Kratzer, L. D. Siracusa, B. A. Quarantillo, R. Evans, and R. M. Liskay.** 1982. Evidence for DNA modification in the maintenance of X-chromosome inactivation of adult mouse tissues. Proc. Natl. Acad. Sci. U.S.A. **79:**5357-5361.

30. **Chapman, V. M., and T. B. Shows.** 1976. Somatic cell genetic evidence for X-chromosome linkage of three enzymes in the mouse. Nature **259:**665-667.

31. **Chung, J. H., M. Whiteley, and G. Felsenfeld.** 1993. A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell **74:**505-14.

32. **Clemson, C. M., J. A. McNeil, H. F. Willard, and J. B. Lawrence.** 1996. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. J. Cell Biol. **132:**259-275.

33. **Cohen, M. M., C. C. Lin, V. Sybert, and E. J. Orecchio.** 1972. Two human X-autosome translocations identified by autoradiography and fluorescence. Am. J. Hum. Genet. **24:**583-597.

34. **Comb, M., and H. M. Goodman.** 1990. CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. Nucleic Acids Res. **18:**3975-3982.

35. **Cooper, G. E., P. L. Bishop, and M. S. Turker.** 1993. Hemidemethylation is sufficient for chromatin relaxation and transcriptional activation of methylated aprt gene in mouse P19 embryonal carcinoma cell line. Somat. Cell Mol. Genet. **19:**221-229.

36. **Cullen, C. R., P. Hubberman, D. C. Kaslow, and B. R. Migeon.** 1986. Comparison of factor IX methylation on human active and inactive X chromosomes: implications for X inactivation and transcription of tissue-specific genes. EMBO. J. **5:**2223-2229.

37. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475-89.

38. **Dracopoli, N. C., W. J. Rettig, A. P. Albino, D. Esposito, N. Archidiacono, M. Rocchi, M. Siniscalco, and L. J. Old.** 1985. Genes controlling gp25/30 cell-surface molecules map to chromosomes X and Y and escape X-inactivation. Am. J. Hum. Genet. **37:**199-207.

39. **Eden, S., and H. Cedar.** 1994. Role of DNA methylation in the regulation of transcription. Curr. Opin. Genet. Dev. **4:**255-259.

40. **Ellis, N., E. Keitges, S. M. Gartler, and M. Rocchi.** 1987. High-frequency reactivation of X-linked genes in Chinese hamster X human hybrid cells. Somat. Cell Mol. Genet. **13:**191-204.

41. **Faisst, S., and S. Meyer.** 1992. Compilation of vertebrate-encoded transcription factors. Nucleic Acids. Res. **20:**3-26.

42. **Fisher, E. M., P. Beer Romero, L. G. Brown, A. Ridley, J. A. McNeil, J. B. Lawrence, H. F. Willard, F. R. Bieber, and D. C. Page.** 1990. Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for Turner syndrome. Cell **63:**1205-1218.

43. **Fried, M., D. M. Crothers, M. M. Garner, and A. Revzin.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic. Acids. Res. **9:**6505-25.

44. **Garner, M. M., A. Revzin, M. M. Garner, and A. Revzin.** 1981. Interaction of catabolite activator protein of Escherichia coli with single-stranded deoxyribonucleic acid. Nucleic Acids. Res. **9:**3047-60.

45. **Gartler, S. M., and B. Burt.** 1964. Replication patterns of bovine sex chromosomes in cell culture. Cytogenetics **3:**135-142.

46. **Gartler, S. M., and A. D. Riggs.** 1983. Mammalian X-chromosome inactivation. Annu. Rev. Genet. **17:**155-190.

47. **Goldman, M. A.** 1988. The chromatin domain as a unit of gene regulation. Bioessays **9:**50-55.

48. **Grant, S. G., and V. M. Chapman.** 1988. Mechanisms of X-chromosome regulation. Annu.Rev.Genet. **22:**199-233.

49. **Graves, J. A., and S. M. Gartler.** 1986. Mammalian X chromosome inactivation: testing the hypothesis of transcriptional control. Somat. Cell Mol. Genet. **12:**275-280.

50. **Green, H., and O. Kehinde.** 1976. Spontaneous heritable changes leading to increased adipose conversion in 3T3 cells. Cell **7:**105-113.

51. **Hahn, S., S. Buratowski, P. A. Sharp, and L. Guarente.** 1989. Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences. Proc. Natl. Acad. Sci. U.S.A. **86:**5718-5722.

52. **Hansen, R. S., N. A. Ellis, and S. M. Gartler.** 1988. Demethylation of specific sites in the 5' region of the inactive X-linked human phosphoglycerate kinase gene correlates with the appearance of nuclease sensitivity and gene expression. Mol. Cell. Biol. **8:**4692-4699.

53. **Hansen, R. S., and S. M. Gartler.** 1990. 5-Azacytidine-induced reactivation of the human X chromosome- linked PGK1 gene is associated with a large region of cytosine demethylation in the 5' CpG island. Proc. Natl. Acad. Sci. U.S.A. **87:**4174-4178.

54. **Harrington, M. A., P. A. Jones, M. Imagawa, and M. Karin.** 1988. Cytosine methylation does not affect binding of transcription factor Sp1. Proc. Natl. Acad. Sci. U.S.A. **85:**2066-2070.

55. **Hebbes, T. R., A. L. Clayton, A. W. Thorne, and C. Crane Robinson.** 1994. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. Embo. J. **13:**1823-30.

56. **Hendrich, B. D.** 1995. Doctrol Thesis. Standford University, Palo Alto.

57. **Holler, M., G. Westin, J. Jiricny, and W. Schaffner.** 1988. Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. Genes Dev. **2:**1127-1135.

58. **Horn, J. M., and A. Ashworth.** 1995. A member of the caudal family of homeobox genes maps to the X- inactivation centre region of the mouse and human X chromosomes. Hum. Mol. Genet. **4:**1041-1047.

59. **Hornstra, I. K.** 1993. Doctoral Thesis. University of Florida, Gainesville.

60. **Hornstra, I. K., D. L. Nelson, S. T. Warren, and T. P. Yang.** 1993. High resolution methylation analysis of the FMR1 gene trinucleotide repeat region in fragile X syndrome. Hum. Mol. Genet. **2:**1659-1665.

61. **Hornstra, I. K., and T. P. Yang.** 1994. High-resolution methylation analysis of the human hypoxanthine phosphoribosyltransferase gene 5' region on the active and inactive X chromosomes: correlation with binding sites for transcription factors. Mol. Cell. Biol. **14:**1419-1430.

62. **Hornstra, I. K., and T. P. Yang.** 1993. In vivo footprinting and genomic sequencing by ligation-mediated PCR. Anal. Biochem. **213:**179-193.

63. **Hornstra, I. K., and T. P. Yang.** 1992. Multiple in vivo footprints are specific to the active allele of the X-linked human hypoxanthine phosphoribosyltransferase gene 5' region: implications for X chromosome inactivation. Mol. Cell. Biol. **12:**5345-5354.

64. **Iguchi Ariga, S. M., and W. Schaffner.** 1989. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. Genes Dev. **3:**612-619.

65. **Jantzen, K., H. P. Fritton, and T. Igo Kemenes.** 1986. The DNase I sensitive domain of the chicken lysozyme gene spans 24 kb. Nucleic. Acids. Res. **14:**6085-99.

66. **Jeppesen, P., and B. M. Turner.** 1993. The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. Cell **74:**281-289.

67. **Johnston, P. G., and B. M. Cattanach.** 1981. Controlling elements in the mouse. IV. Evidence of non-random X- inactivation. Genet. Res. **37:**151-160.

68. **Kalos, M., and R. E. Fournier.** 1995. Position-independent transgene expression mediated by boundary elements from the apolipoprotein B chromatin domain. Mol. Cell. Biol. **15:**198-207.

69. **Kay, G. F., S. C. Barton, M. A. Surani, and S. Rastan.** 1994. Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. Cell **77:**639-650.

70. **Keitges, E., and S. M. Gartler.** 1986. Dosage of the Sts gene in the mouse. . Am. J. Hum. Genet. **39:**470-476.

71. **Keith, D. H., J. Singer Sam, and A. D. Riggs.** 1986. Active X chromosome DNA is unmethylated at eight CCGG sites clustered in a guanine-plus-cytosine-rich island at the 5' end of the gene for phosphoglycerate kinase. Mol. Cell Biol. **6:**4122-4125.

72. **Kerem, B. S., R. Goitein, C. Richler, M. Marcus, and H. Cedar.** 1983. In situ nick-translation distinguishes between active and inactive X chromosomes. Nature **304:**88-90.

73. **Keshet, I., J. Lieman Hurwitz, and H. Cedar.** 1986. DNA methylation affects the formation of active chromatin. Cell **44:**535-543.

74. **Kierszenbaum, A. L., and L. L. Tres.** 1974. Nucleolar and perichromosomal RNA synthesis during meiotic prophase in the mouse testis. J. Cell Biol. **60:**39-53.

75. **Kierszenbaum, A. L., and L. L. Tres.** 1974. Transcription sites in spread meiotic prophase chromosomes from mouse spermatocytes. J. Cell Biol. **63:**923-935.

76. **Kofman Alfaro, S., and A. C. Chandley.** 1970. Meiosis in the male mouse. An autoradiographic investigation. Chromosoma **31:**404-420.

77. **Kramer, J. M., and R. P. Erickson.** 1981. Developmental program of PGK-1 and PGK-2 isozymes in spermatogenic cells of the mouse: specific activities and rates of synthesis. Dev. Biol. **87:**37-45.

78. **Kulozik, A. E., S. Bail, A. Bellan Koch, C. R. Bartram, E. Kohne, and E. Kleihauer.** 1991. The proximal element of the beta globin locus control region is not functionally required in vivo. J. Clin. Invest. **87:**2142-6.

79. **Lawson, G. M., B. J. Knoll, C. J. March, S. L. Woo, M. J. Tsai, and O. M. BW.** 1982. Definition of 5' and 3' structural boundaries of the chromatin domain containing the ovalbumin multigene family. J. Biol. Chem. **257:**1501-7.

80. **Lee, J. T., W. M. Strauss, J. A. Dausman, and R. Jaenisch.** 1996. A 450 kb transgene displays properties of the mammalian X-inactivation center. Cell **86:**83-94.

81. **Lee, M. S., and W. T. Garrard.** 1992. Uncoupling gene activity from chromatin structure: promoter mutations can inactivate transcription of the yeast HSP82 gene without eliminating nucleosome-free regions. Proc. Natl. Acad. Sci. U.S.A. **89:**9166-9170.

82. **Levine, A., G. L. Cantoni, and A. Razin.** 1992. Methylation in the preinitiation domain suppresses gene transcription by an indirect mechanism. Proc. Natl. Acad. Sci. U.S.A. **89:**10119-10123.

83. **Levy Wilson, B., and C. Fortier.** 1989. The limits of the DNase I-sensitive domain of the human apolipoprotein B gene coincide with the locations of chromosomal anchorage loops and define the 5' and 3' boundaries of the gene. J Biol. Chem. **264:**21196-204.

84. **Lewis, J. D., R. R. Meehan, W. J. Henzel, I. Maurer Fogy, P. Jeppesen, F. Klein, and A. Bird.** 1992. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell **69:**905-914.

85. **Lin, D., and A. C. Chinault.** 1988. Comparative study of DNase I sensitivity at the X-linked human HPRT locus. Somat. Cell Mol. Genet. **14:**261-272.

86. **Liskay, R. M., and R. J. Evans.** 1980. Inactive X chromosome DNA does not function in DNA-mediated cell transformation for the hypoxanthine phosphoribosyltransferase gene. Proc. Natl. Acad. Sci. U.S.A. **77:**4895-4898.

87. **Litt, M. D., I. K. Hornstra, and T. P. Yang.** 1996. In vivo footprinting and high-resolution methylation analysis of the mouse hypoxanthine and phosphoribosyltransferase gene 5' region on the active and inactive X chromosomes. Mol. Cell. Biol. **16:**6190-6199.

88. **Loc, P. V., and W. H. Stratling.** 1988. The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. Embo. J **7:**655-64.

89. **Lock, L. F., D. W. Melton, C. T. Caskey, and G. R. Martin.** 1986. Methylation of the mouse hprt gene differs on the active and inactive X chromosomes. Mol. Cell. Biol. **6:**914-924.

90. **Lyon, M. F.** 1961. Gene action in the X chromosome of the mouse (Mus musculus L.). Nature **190:**372.

91. **Maniatis, T., S. Goodbourn, and J. A. Fischer.** 1987. Regulation of inducible and tissue-specific gene expression. Science **236:**1237-1245.

92. **Martinez Balbas, M. A., A. Dey, S. K. Rabindran, K. Ozato, and C. Wu.** 1995. Displacement of sequence-specific transcription factors from mitotic chromatin. Cell **83:**29-38.

93. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65:**499-560.

94. **McBurney, M. W.** 1988. X chromosome inactivation: a hypothesis. Bioessays **9:**85-88.

95. **McCarrey, J. R., W. M. Berg, S. J. Paragioudakis, P. L. Zhang, D. D. Dilworth, B. L. Arnold, and J. J. Rossi.** 1992. Differential transcription of Pgk genes during spermatogenesis in the mouse. Dev. Biol. **154:**160-168.

96. **McCarrey, J. R., and D. D. Dilworth.** 1992. Expression of Xist in mouse germ cells correlates with X- chromosome inactivation. Nat. Genet. **2:**200-203.

97. **McCarrey, J. R., D. D. Dilworth, and R. M. Sharp.** 1992. Semiquantitative analysis of X-linked gene expression during spermatogenesis in the mouse: ethidium-bromide staining of RT-PCR products. Genet. Anal. Tech. Appl. **9:**117-123.

98. **McKnight, R. A., A. Shamay, L. Sankaran, R. J. Wall, and L. Hennighausen.** 1992. Matrix-attachment regions can impart position-independent regulation of a tissue-specific gene in transgenic mice. Proc. Natl. Acad. Sci. U.S.A. **89:**6943-7.

99. **Means, A. L., and P. J. Farnham.** 1990. Transcription initiation from the dihydrofolate reductase promoter is positioned by HIP1 binding at the initiation site. Mol. Cell. Biol. **10:**653-661.

100. **Meehan, R. R., J. D. Lewis, S. McKay, E. L. Kleiner, and A. P. Bird.** 1989. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. Cell **58:**499-507.

101. **Melton, D. W., D. S. Konecki, J. Brennand, and C. T. Caskey.** 1984. Structure, expression, and mutation of the hypoxanthine phosphoribosyltransferase gene. Proc. Natl. Acad. Sci. U.S.A. **81:**2147-2151.

102. **Melton, D. W., C. McEwan, A. B. McKie, and A. M. Reid.** 1986. Expression of the mouse HPRT gene: deletional analysis of the promoter region of an X-chromosome linked housekeeping gene. Cell **44:**319-328.

103. **Migeon, B. R.** 1994. X-chromosome inactivation: molecular mechanisms and genetic consequences. Trends. Genet. **10:**230-235.

104. **Migeon, B. R., L. J. Shapiro, R. A. Norum, T. Mohandas, J. Axelman, and R. L. Dabora.** 1982. Differential expression of steroid sulphatase locus on active and inactive human X chromosome. Nature **299:**838-840.

105. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence- specific DNA binding proteins. Science **245:**371-378.

106. **Mohandas, T., R. L. Geller, P. H. Yen, J. Rosendorff, R. Bernstein, A. Yoshida, and L. J. Shapiro.** 1987. Cytogenetic and molecular studies on a recombinant human X chromosome: implications for the spreading of X chromosome inactivation. Proc. Natl. Acad. Sci. U.S.A. **84:**4954-4958.

107. **Mohandas, T., R. S. Sparkes, and L. J. Shapiro.** 1981. Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. Science **211:**393-396.

108. **Monesi, V.** 1971. Chromosome activities during meiosis and spermiogenesis. J. Reprod. Fertil. Suppl. **13:**Suppl 13:1-14.

109. **Naumova, A. K., R. M. Plenge, L. M. Bird, M. Leppert, K. Morgan, H. F. Willard, and C. Sapienza.** 1996. Heritability of X chromosome--inactivation phenotype in a large family [see comments]. Am. J. Hum. Genet. **58:**1111-1119.

110. **Nussbaum, R. L., S. D. Airhart, and D. H. Ledbetter.** 1983. Expression of the fragile (X) chromosome in an interspecific somatic cell hybrid. Hum. Genet. **64:**148-150.

111. **Odartchenko, N., and M. Pavillard.** 1970. Late DNA replication in male mouse meiotic chromosomes. Science **167:**1133-1134.

112. **Ohno, S., W. D. Kaplan, and R. Kinosita.** 1959. Formation of sex chromatin by a single-X chromosome in liver cells of Rattu norvegicus. Exp. Cell Res. **18:**415-418.

113. **Panning, B., and R. Jaenisch.** 1996. DNA hypomethylation can activate Xist expression and silence X- linked genes. Genes Dev. **10:**1991-2002.

114. **Park, J. G., and V. M. Chapman.** 1994. CpG island promoter region methylation patterns of the inactive- X-chromosome hypoxanthine phosphoribosyltransferase (Hprt) gene. Mol. Cell. Biol. **14:**7975-7983.

115. **Patel, P. I., P. E. Framson, C. T. Caskey, and A. C. Chinault.** 1986. Fine structure of the human hypoxanthine phosphoribosyltransferase gene. Mol.Cell. Biol. **6:**393-403.

116. **Patel, P. I., R. L. Nussbaum, P. E. Gramson, D. H. Ledbetter, C. T. Caskey, and A. C. Chinault.** 1986. Organization of the HPRT gene and related sequences in the human genome. Somat Cell Mol Genet **10:**483-93.

117. **Pathak, S., and F. F. Elder.** 1980. Silver-stained accessory structures on human sex chromosomes. Hum. Genet. **54:**171-175.

118. **Penny, G. D., G. F. Kay, S. A. Sheardown, S. Rastan, and N. Brockdorff.** 1996. Requirement for Xist in X chromosome inactivation [see comments]. Nature **379:**131-137.

119. **Pfeifer, G. P., and A. D. Riggs.** 1991. Chromatin differences between active and inactive X chromosomes revealed by genomic footprinting of permeabilized cells using DNase I and ligation-mediated PCR. Genes Dev. **5:**1102-1113.

120. **Pfeifer, G. P., S. D. Steigerwald, R. S. Hansen, S. M. Gartler, and A. D. Riggs.** 1990. Polymerase chain reaction-aided genomic sequencing of an X chromosome-linked CpG island: methylation patterns suggest clonal inheritance, CpG site autonomy, and an explanation of activity state stability. Proc. Natl. Acad. Sci. U.S.A. **87:**8252-8256.

121. **Pfeifer, G. P., S. D. Steigerwald, R. S. Hansen, S. M. Gartler, and A. D. Riggs.** 1990. Polymerase chain reaction-aided genomic sequencing of an X chromosome-linked CpG island: methylation patterns suggest clonal inheritance, CpG site autonomy, and an explanation of activity state stability. Proc. Natl. Acad. Sci. U.S.A. **87:**8252-8256.

122. **Pfeifer, G. P., R. L. Tanguay, S. D. Steigerwald, and A. D. Riggs.** 1990. In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. Genes Dev. **4:**1277-1287.

123. **Rastan, S.** 1983. Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre. J. Embryol. Exp. Morphol. **78:**1-22.

124. **Rastan, S., and E. J. Robertson.** 1985. X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. J. Embryol. Exp. Morphol. **90:**379-388.

125. **Reitman, M., E. Lee, H. Westphal, and G. Felsenfeld.** 1993. An enhancer/locus control region is not sufficient to open chromatin. Mol. Cell. Biol. **13:**3990-8.

126. **Riley, D. E., T. K. Canfield, and S. M. Gartler.** 1984. Chromatin structure of active and inactive human X chromosomes. Nucleic. Acids. Res. **12:**1829-1845.

127. **Riley, D. E., M. A. Goldman, and S. M. Gartler.** 1986. Chromatin structure of active and inactive human X-linked phosphoglycerate kinase gene. Somat. Cell Mol. Genet. **12:**73-80.

128. **Rincon-Limas, D. E., F. Amaya-Manzanares, M. L. Nino-Rosales, Y. Yu, T. P. Yang, and P. I. Patel.** 1995. Ubiquitous and Neuronal DNA-Binding Proteins Interact with a Negative Regulatory Element of the Human Hypoxanthine Phosphoribosyltransferase Gene. Mol. Cell. Biol. **15:**6561-6571.

129. **Romrell, L. J., and M. G. O'Rand.** 1978. Capping and ultrastructural localization of sperm surface isoantigens during spermatogenesis. Dev. Biol. **63:**76-93.

130. **Rougeulle, C., and P. Avner.** 1996. Cloning and characterization of a murine brain specific gene Bpx and its human homologue lying within the Xic candidate region. Hum. Mol. Genet. **5:**41-49.

131. **Salido, E. C., P. H. Yen, T. K. Mohandas, and L. J. Shapiro.** 1992. Expression of the X-inactivation-associated gene XIST during spermatogenesis. Nat. Genet. **2:**196-199.

132. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular Cloning: A Laboratory Manual 2nd ed, 2 ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

133. **Sasaki, T., R. S. Hansen, and S. M. Gartler.** 1992. Hemimethylation and hypersensitivity are early events in transcriptional reactivation of human inactive X-linked genes in a hamster x human somatic cell hybrid. Mol.Cell. Biol. **12:**3819-3826.

134. **Seto, E., Y. Shi, and T. Shenk.** 1991. YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro. Nature **354:**241-245.

135. **Singer Sam, J., and A. D. Riggs.** 1993. X chromosome inactivation and DNA methylation. EXS. **64:**358-384.

136. **Singer Sam, J., M. O. Robinson, A. R. Bellve, M. I. Simon, and A. D. Riggs.** 1990. Measurement by quantitative PCR of changes in HPRT, PGK-1, PGK-2, APRT, MTase, and Zfy gene transcripts during mouse spermatogenesis. Nucleic. Acids. Res. **18:**1255-1259.

137. **Slansky, J. E., Y. Li, W. G. Kaelin, and P. J. Farnham.** 1993. A protein synthesis-dependent increase in E2F1 mRNA correlates with growth regulation of the dihydrofolate reductase promoter [published erratum appears in Mol Cell Biol 1993 Nov;13(11):7201]. Mol. Cell Biol. **13:**1610-1618.

138. **Smale, S. T., and D. Baltimore.** 1989. The "initiator" as a transcription control element. Cell **57:**103-113.

139. **Stief, A., D. M. Winter, W. H. Stratling, and A. E. Sippel.** 1989. A nuclear DNA attachment element mediates elevated and position-independent gene activity. Nature **341:**343-5.

140. **Stout, J. T., and C. T. Caskey.** 1985. HPRT: gene structure, expression, and mutation. Annu. Rev. Genet. **19:**127-148.

141. **Takagi, N.** 1974. Differentiation of X chromosomes in early female mouse embryos. Exp. Cell Res. **86:**127-135.

142. **Therman, E., G. E. Sarto, and K. Patau.** 1974. Center for Barr body condensation on the proximal part of the human Xq: a hypothesis. Chromosoma **44:**361-366.

143. **Tommasi, S., J. M. LeBon, A. D. Riggs, and J. Singer Sam.** 1993. Methylation analysis by genomic sequencing of 5' region of mouse Pgk-1 gene and a cautionary note concerning the method. Somat. Cell Mol. Genet. **19:**529-541.

144. **Toniolo, D., G. Martini, B. R. Migeon, and R. Dono.** 1988. Expression of the G6PD locus on the human X chromosome is associated with demethylation of three CpG islands within 100 kb of DNA. EMBO J. **7:**401-406.

145. **Turner, B. M.** 1991. Histone acetylation and control of gene expression. J. Cell Sci. **99:**13-20.

146. **VandeBerg, J. L., D. W. Cooper, and P. J. Close.** 1976. Testis specific phosphoglycerate kinase B in mouse. J. Exp. Zool. **198:**231-240.

147. **Venolia, L., and S. M. Gartler.** 1983. Comparison of transformation efficiency of human active and inactive X-chromosomal DNA. Nature **302:**82-83.

148. **Venolia, L., S. M. Gartler, E. R. Wassman, P. Yen, T. Mohandas, and L. J. Shapiro.** 1982. Transformation with DNA from 5-azacytidine-reactivated X chromosomes. Proc. Natl. Acad. Sci. U.S.A. **79:**2352-2354.

149. **Weintraub, H., A. Larsen, and M. Groudine.** 1981. Alpha-Globin-gene switching during the development of chicken embryos: expression and chromosome structure. Cell **24:**333-44.

150. **Weis, L., and D. Reinberg.** 1992. Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. FASEB J. **6:**3300-3309.

151. **Wolf, S. F., D. J. Jolly, K. D. Lunnen, T. Friedmann, and B. R. Migeon.** 1984. Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X chromosome: implications for X-chromosome inactivation. Proc. Natl. Acad. Sci.U.S.A. **81:**2806-2810.

152. **Wolf, S. F., and B. R. Migeon.** 1985. Clusters of CpG dinucleotides implicated by nuclease hypersensitivity as control elements of housekeeping genes. Nature **314:**467-469.

153. **Wolf, S. F., and B. R. Migeon.** 1982. Studies of X chromosome DNA methylation in normal human cells. Nature **295:**667-671.

154. **Wu, J., E. C. Salido, P. H. Yen, T. K. Mohandas, H. H. Heng, L. C. Tsui, J. Park, V. M. Chapman, and L. J. Shapiro.** 1994. The murine Xe169 gene escapes X-inactivation like its human homologue. Nat. Genet. **7:**491-496.

155. **Yang, T. P., and C. T. Caskey.** 1987. Nuclease sensitivity of the mouse HPRT gene promoter region: differential sensitivity on the active and inactive X chromosomes. Mol. Cell. Biol. **7:**2994-2998.

156. **Yen, P. H., P. Patel, A. C. Chinault, T. Mohandas, and L. J. Shapiro.** 1984. Differential methylation of hypoxanthine phosphoribosyltransferase genes on active and inactive human X chromosomes. Proc.Natl.Acad.Sci.U.S.A. **81:**1759-1763.

157. **Zinn, A. R., S. L. Bressler, P. Beer Romero, D. A. Adler, V. M. Chapman, D. C. Page, and C. M. Disteche.** 1991. Inactivation of the Rps4 gene on the mouse X chromosome [published erratum appears in Genomics 1992 Jul;13(3):915]. Genomics **11:**1097-1101.

BIOGRAPHICAL SKETCH

I was born October 7, 1965, in Los Angles, California. I spent most of my childhood in Kelseyville, a small town nestled under Mt. Konocti in Northern California. My father, Marvin D. Litt, worked for the phone company while I lived in Kelseyville. Presently, my father is working for NTI and appears to be having the time of his life traveling from one exotic location to another. My mother, Donna L. Litt, worked around the house while doing other various part time jobs. She is also having fun traveling around the world. I attended and graduated from Kelseyville High School in 1984. I continued my education at the University of California at Davis, studying human physiology. At Davis, I worked part time as a Laboratory Assistant in Dr. Carlson's lab. This was my first exposure to the new and growing science of molecular biology. I also met my wife, Christine S. Marino, at Davis while attending a rock concert. We have been together since that day. Today, I am continuing to study molecular biology. After I finish here, I will be working as a postdoctoral fellow in Gary Felsenfeld's Lab at the NIH studying chromatin structure.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Thomas P. Yang, Chairman
Associate Professor of
Biochemistry and Molecular
Biology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Thomas C. Rowe
Associate Professor of
Pharmacology and
Therapeutics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert J. Ferl
Professor of Horticultural
Science

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Peter M. McGuire

Peter M. McGuire
Assistant Professor of
Biochemistry and Molecular
Biology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Philip J. Laipis

Philip J. Laipis
Professor of Biochemistry and
Molecular Biology

This dissertation was submitted to the Graduate Faculty of the College of Medicine and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1997

Mohan K Raizada

Dean, College of Medicine

Karen A. Holbrook

Dean, Graduate School

UNIVERSITY OF FLORIDA
3 1262 08557 0355